

Projekt pn. „*Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych*”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

# **Wybrane metody analizy strukturalnej związków małowcząsteczkowych i biomakrocząsteczek. Od biologii, przez fizykę, do chemii.**

Prof. dr hab. Andrzej Wojtczak  
Dr Anna Kozakiewicz

UMK Toruń 2014

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## **Część I**

Prof. dr hab. Andrzej Wojtczak

1. Podstawowe pojęcia w krytalografii	3
2. Dyfrakcja promieniowania rentgenowskiego na kryształach	13
3. Teoria czynników struktury i problem fazowy	19
4. Metody małocząsteczkowe rozwiązywania problemu fazowego	23
5. Metody proteomiki strukturalnej	28
6. Elementy bioinformatyki – modelowanie molekularne	34
7. Zastosowanie metod analizy strukturalnej	39
8. Zastosowanie metod analizy strukturalnej cd.	39
9. Literatura do części I	46

## **Część II**

Dr Anna Kozakiewicz

1. Projektowanie leków - krytalografia jako narzędzie w chemii, biochemii i medycynie	47
2. Otrzymywanie i oczyszczanie białek na potrzeby krytalografii	54
3. Krytalizacja białek	62
4. Pomiar synchrotronowy	68
5. Rozwiązywanie struktur krytalicznych	73
6. Walidacja i deponowanie struktury	77
7. Wizualizacja struktur białkowych oraz małocząsteczkowych	82
8. Dokowanie ligandów do receptora	87
9. Literatura do części II	92

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Część I

### Wykład 1: Podstawowe pojęcia w krytalografii

Krytalografia jest interdyscyplinarną dziedziną wiedzy, mającą swoje źródło w mineralogii. Ze względu na opis materii, zjawisk związanych z badaniami oraz naturę obiektów i problemy dające się rozwiązać, krytalografia jest nauką interdyscyplinarną na pograniczu fizyki (metody badawcze) matematyki (opis zjawisk), chemii i biologii (związek budowy przestrzennej z właściwościami i funkcją cząsteczek i biocząsteczek).

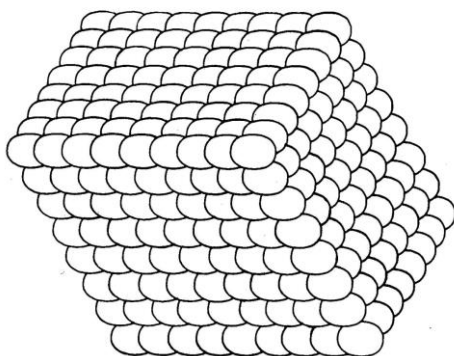
Podwaliny współczesnej krytalografii stworzono w XVII wieku. Wśród ważnych podstaw jest sformułowane przez Nicolausa Steno prawo stałości kątów mówiące, że kąty pomiędzy ścianami kryształów danej substancji otrzymanymi w tych samych warunkach fizykochemicznych (np. pH, temperatura, rozpuszczalnik) są stałe. Racjonalne podstawy tego prawa zaproponował mineralog Rene Haüy, stawiając hipotezę, że kryształy są zbudowane z identycznych fragmentów powielanych w przestrzeni. Rozważmy proces wzrostu kryształu zbudowanego z takich samych cząsteczek. Ze względu na budowę chemiczną, każda cząsteczka ma preferowane kierunki oddziaływań z cząsteczkami sąsiadującymi. Energia oddziaływań międzycząsteczkowych zależy od fragmentów cząsteczek (grup funkcyjnych) w nich uczestniczących. Jeśli pozwolimy, aby układ, w którym następuje wzrost kryształów osiągnął równowagę termodynamiczną, możemy oczekiwać, że każda cząsteczka będzie tworzyła analogiczne oddziaływania z otoczeniem, a odległości pomiędzy cząsteczkami w jednakowych kierunkach będą identyczne (**Rys. 1.1**). W przestrzeni można wybrać trzy kierunki odpowiadające największej energii oddziaływań między cząsteczkami jako kierunki osi układu współrzędnych, w którym dokonamy opisu kryształu. Jeśli otoczenie każdej cząsteczki w kryształach jest jednakowe, to kryształ ma budowę periodyczną. Jednostkowe translacje w kierunkach wybranych osi układu (wersowy osi) odpowiadają odległościom pomiędzy sąsiednimi cząsteczkami. Wybór osi układu może być dokonany przez analizę morfologii kryształu – kierunki o największej energii oddziaływań międzycząsteczkowych będą pojawiały się w pokroju kryształu jako najdłuższe krawędzie (**Rys. 1.2**) odpowiadające kierunkom o największej liniowej gęstości materii, zaś największe ściany odpowiadają płaszczyznom o największej gęstości powierzchniowej, co także wiąże się z optymalnymi termodynamicznie oddziaływaniami. Jeśli cząsteczki wykazują symetrię w budowie przestrzennej, to sieć oddziaływań przez nie tworzonych może także wykazać taką symetrię, co znajdzie odbicie w symetrii budowy kryształu.

W oparciu o powyższe można zdefiniować kryształ jako jednorodnie chemicznie ciało stałe o uporządkowanej budowie wewnętrznej (periodyczność, symetria rozkładu materii), w którym właściwości skalarne są jednakowe w każdym kierunku, a właściwości wektorowe są jednakowe w kierunkach równoległych lub związanych symetrią.

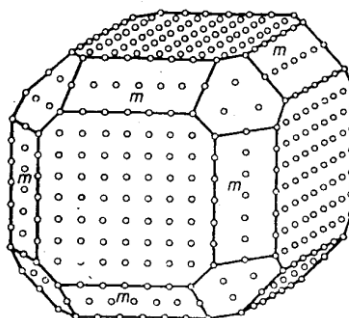
Opisu kryształu można dokonać podając współrzędne wszystkich atomów w kryształach. Wtedy mówi się o sieci krystalicznej. Można jednak uprościć opis, zastępując cząsteczki (np. białka, wirusy, asocjaty cząsteczek) przez punkty zwane węzłami. Taki opis jest znacznie prostszy, ale

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

zachowuje wiernie zależności geometryczne między składnikami sieci, w tym odległości i symetrię sieci. Przy takim opisie posługujemy się pojęciem sieci przestrzennej.



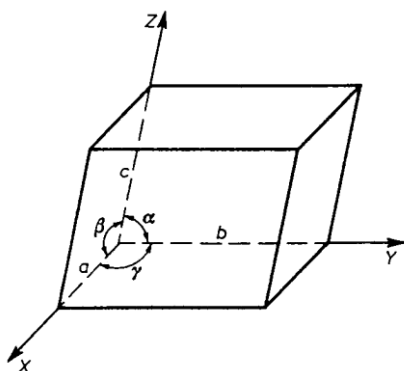
Rys. 1.1.



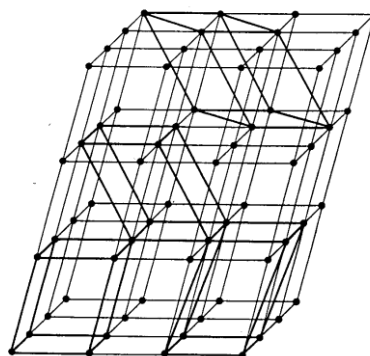
Rys. 1.2

Na osie układu wybieramy trzy niewspółpłaszczyznowe proste, tworzące prawoskrętny układ odniesienia. Dla opisu osi układu i translacji jednostkowych przyjęto konwencję opisaną na **Rys. 1.3**. Wersorami osi X, Y i Z są odpowiednio wektory  $a_0$ ,  $b_0$  i  $c_0$ . Jeśli za początek układu przyjmiemy jeden z identycznych węzłów, to jednostkowe przesunięcia  $a_0$ ,  $b_0$  i  $c_0$  oraz ich sumy zdefiniują położenia ośmiu węzłów (reprezentacji cząsteczek) w narożach równoległościanu. Taki równoległościan nazywa się komórką elementarną, i odpowiada on podstawowej jednostce budowy kryształu postulowanej przez Rene Haüy. Kształt komórki zależy od wektorów translacji między węzłami. Długości wersorów osi odpowiadają długości krawędzi komórki elementarnej, a kąty między krawędziami odpowiadają kątom między osiami układu. Długości wersorów osi i kąty między osiami ( $a_0$ ,  $b_0$ ,  $c_0$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ), nazywa się parametrami komórki elementarnej. Wyboru komórki elementarnej w periodycznej sieci można dokonać na wiele sposobów (**Rys. 1.4**). Wybór musi jednak spełniać kilka warunków:

- Zawartość komórki elementarnej jest reprezentatywna dla całego kryształu
- Wybór musi odpowiadać minimalnej objętości komórki, jak największej liczbie kątów prostych między osiami i najwyższej symetrii.



Rys. 1.3.



Rys. 1.4

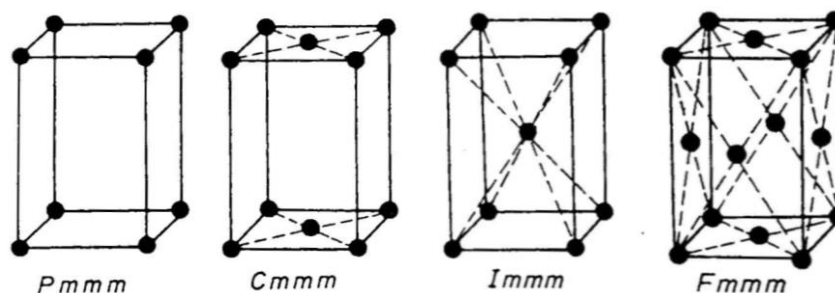
Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Ze względu na parametry komórki elementarnej, kryształy zalicza się do jednego z 7 układów krystalograficznych:

Układ trójskośny	$a \neq b \neq c \quad \alpha \neq \beta \neq \gamma$	L1
Układ jednoskośny	$a \neq b \neq c \quad \alpha = \gamma = 90 < \beta$	L2
Układ rombowy	$a \neq b \neq c \quad \alpha = \beta = \gamma = 90$	3L2
Układ tetragonalny	$a = b \neq c \quad \alpha = \beta = \gamma = 90$	L4
Układ trygonalny	$a = b = c \quad \alpha = \beta = \gamma \neq 90$ (romboedryczna)	L3
	$a = b \neq c \quad \alpha = \beta = 90 \quad \gamma = 120$ (heksagonalna)	
Układ heksagonalny	$a = b \neq c \quad \alpha = \beta = 90 \quad \gamma = 120$	L6
Układ regularny	$a = b = c \quad \alpha = \beta = \gamma = 90$	4L3

Zestawienie zawiera też osie symetrii charakterystyczne dla danego układu.

W rzeczywistej periodycznej sieci krystalicznej cząsteczki leżą nie tylko wokół naroży komórki, ale także na środkach ścian i w środku przestrzennym. Cztery podstawowe typy komórek: prymitywną P, o centrowanych podstawach C, wewnątrzcentrowaną I oraz ściennie centrowaną F przedstawia **Rys. 1.5**. Matematyk francuski Auguste Bravais wyprowadził wszystkie 14 możliwych typów sieci translacyjnych, uwzględniających 4 typy centrowania i 7 układów krystalograficznych. Dla pokazanych typów komórek, położenie węzłów opisane jest liniową kombinacją jednostkowych translacji  $(a_0, b_0, c_0)$  w komórce P, oraz dodatkowo translacjami  $(a+b)/2$  i  $(a+b)/2 + c$  w komórce C,  $(a+b+c)/2$  w komórce I oraz  $(a+b)/2$ ;  $(a+b)/2 + c$ ;  $(a+c)/2$ ;  $(a+c)/2 + b$ ;  $(b+c)/2$ ;  $(b+c)/2 + a$  w komórce typu F. Taki rozkład węzłów spełnia wymogi grupy przemiennej translacji w przestrzeni.



**Rys. 1.5.**

Grupa translacyjna to zbiór elementów grupy – translacji oraz działanie grupowe – suma wektorów translacji, określone w przestrzeni na przykład trójwymiarowej. Zbiór translacji stanowiący grupę przemianą musi spełniać pięć warunków:

- Złożenie dwóch translacji daje translację należącą do grupy
- Łączność. Dla każdych translacji  $a, b, c \in G \quad (a+b)+c = a+(b+c)$
- Element neutralny  $e \in G$  taki, że dla każdego  $a \in G \quad a+e = e+a = a$
- El. przeciwny. Dla każdego  $a \in G$  istnieje  $a^{-1} \in G$  takie, że  $a+a^{-1} = a^{-1}+a = e$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

- Przemienność. Dla każdych  $a, b \in G$  zachodzi  $a+b = b+a$

Zakładając sferyczną symetrię węzłów oraz ich położenie w narożach komórki elementarnej, łatwo wyliczyć, że  $\frac{1}{2}$  węzła na ścianie należy do wnętrza komórki elementarnej. W przypadku węzłów w narożach komórki prostokątnej,  $\frac{1}{8}$  każdego z nich należy do wnętrza komórki. Łatwo wykazać, że dla komórek o dowolnym kształcie suma wkładów takich węzłów daje 1 węzeł wewnątrz danej komórki. Dlatego dla komórki P, C, I oraz F liczba węzłów Z w objętości komórki wynosi odpowiednio 1, 2, 2 oraz 4.

Opis zawartości kryształu upraszcza się do opisu zawartości pojedynczej komórki elementarnej oraz podania wektorów jednostkowych translacji  $a_0, b_0, c_0$ , których liniowa kombinacja definiuje położenie kolejnych komórek w sieci krystalicznej. To nasuwa konieczność zastosowania geometrii analitycznej lub wprowadzenia odpowiedniej symboliki stosowanej w krystalografii. Położenie punktu (np. położenie równowagowe atomu) definiuje się przez podanie współrzędnych ułamkowych  $x, y, z$ . Są to współrzędne względne, liczone jako ułamki wektorów jednostkowych translacji wzdłuż osi X Y Z układu. Dlatego położenie każdego punktu  $x, y, z$  można opisać wektorem z punktem zaczepienia w początku układu:

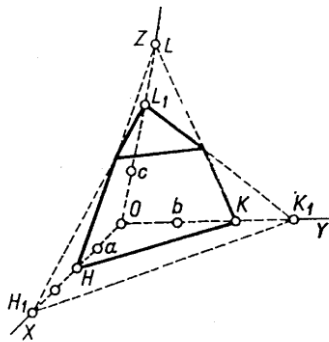
$$\vec{r} = x\vec{a} + y\vec{b} + z\vec{c}$$

Z powyższego wynika, że współrzędne punktów należących do komórki elementarnej przyjmują wartości z przedziału  $[0,1]$ .

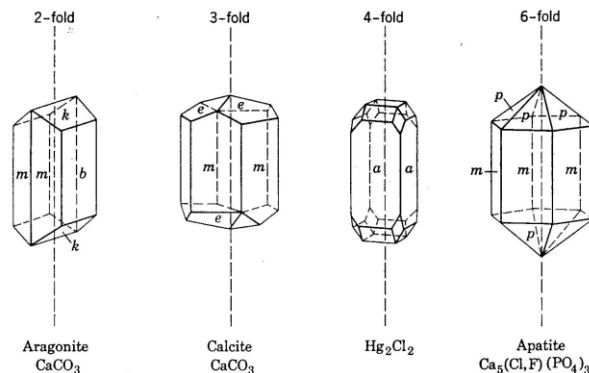
Proste i płaszczyzny w sieci przestrzennej definiowane są przez zbiory węzłów (**Rys. 1.2 i 1.4**). Kierunek prostej w przestrzeni można opisać przez określenie dwóch punktów  $x_1, y_1, z_1$  oraz  $x_2, y_2, z_2$  należących do tej prostej. Wskaźniki w symbolu prostej  $[uvw]$  oblicza się jako różnicę współrzędnych pary punktów  $[\Delta x, \Delta y, \Delta z]$ . Przykładowo para punktów 112 i 223 definiuje prostą  $[111]$ . Jeżeli rozważymy osie X Y i Z, przechodzące przez początek układu oraz punkty na końcach wektorów  $a_0, b_0, c_0$ , to osie mają symbole X  $[100]$  Y  $[010]$  i Z  $[001]$ . Należy zauważyć, że wszystkie proste równoległe w przestrzeni mają identyczny symbol  $[uvw]$ , co jest zgodne z przyjętą definicją kryształu.

Płaszczyzny w przestrzeni można jednoznacznie określić przez podanie współrzędnych 3 punktów. Aby wprowadzić symbole płaszczyzn trzeba rozważyć parę płaszczyzn przecinających osie układu w punktach HKL oraz  $H_1, K_1, L_1$ , odpowiadających węzłom na osiach (**Rys. 1.6**). Położenie tych punktów jest opisane całkowitą wielokrotnością wektorów osi  $a_0, b_0, c_0$ . To prowadzi do prawa wymiernych stosunków odcinków: ilorazy długości odcinków odcinanych na osiach przez dwie płaszczyzny są liczbami wymiernymi  $OH/OH_1, OK/OK_1, OL/OL_1 \in W$ .

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki



Rys. 1.6.



Rys 1.7

Na tej podstawie angielski krystalograf Miller wprowadził symbolikę płaszczyzn. Jeśli założyć, że odcinki OH, OK. i OL odpowiadają odcinkom jednostkowym  $OH=a_0$ ,  $OK.=b_0$ ,  $OL=c_0$ , to ściana przecinająca osie w punktach H K i L może być ścianą odniesienia, a położenie każdej innej płaszczyzny można opisać względem tej płaszczyzny. Przekształcając trzy liczby wymierne do liczb całkowitych uzyskujemy wskaźniki w symbolu Millera płaszczyzny (hkl):

$$h = a_0/OH_1 \quad k = b_0/OK_1 \quad l = c_0/OL_1 \quad \text{gdzie } h, k, l \in \mathbb{C}$$

Z definicji kryształu wynika, że płaszczyzny równoległe w przestrzeni są sobie równoważne, gdyż przecinają osie układu w kolejnych węzłach na osiach. Można zatem opisać w sposób kompletny komórkę elementarną i scharakteryzować jednoznacznie układ krystalograficzny przy pomocy tylko 4 płaszczyzn. Trzy z nich są definiowane przez pary osi układu XY, XZ i YZ, a czwarta przecina osie w punktach określających długość wektorów osi (**Rys. 1.3**). Ten zbiór ścian nazywa się czworościanem zasadniczym.

Równoległość do osi układu oznacza nieskończoną długość odcinka odcinanego na tej osi. Z definicji symboliki Millera wskaźnik obliczony np. dla osi X jako  $\lim_{a_0 \rightarrow \infty} (a_0/\infty) = 0$ . Dlatego 3 ściany czworościanu odpowiadające ścianom komórki elementarnej mają symbole (100) (010) i (001), a czwarta ściana, wyznaczająca długość wektorów osi ma symbol (111).

W krystalografii używa się pojęcia pasa – zbioru płaszczyzn równoległych do wspólnej prostej, zwanej osią pasa. Przynależność do pasa prostej [uvw], czyli równoległość płaszczyzny (hkl) do tej prostej opisuje równanie pasa:

$$hx + ky + lz = 0$$

Niespełnienie tego warunku oznacza, że prosta [uvw] i płaszczyzna (hkl) nie są do siebie równoległe. Konsekwencją równania pasa jest, że wszystkie płaszczyzny (0kl) spełniają równanie pasa dla osi X [100], zaś płaszczyzny (h0l) i (hk0) są równoległe do osi Y i Z.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

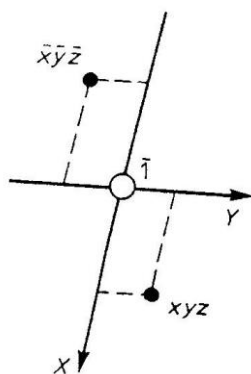
Rozkład materii w sieci krystalicznej wykazuje symetrię widoczną w morfologii kryształu na **Rys. 1.7**. Symetrie tak jak translacje są przekształceniami izometrycznymi – zachowującymi odległości. Do opisu symetrii **S** stosuje się operatory macierzowe:

$$\mathbf{S} \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \text{ działające na punkty } \mathbf{X} \text{ dając obraz } \mathbf{X}'$$

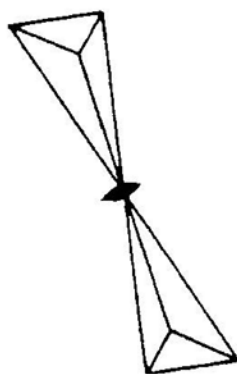
Postać macierzy jest łatwa do otrzymania. Rozważenie iloczynu macierzy **S** i macierzy kolumnowych opisujących końce wektorów  $a_0, b_0, c_0$ . Przykładowo:

$$\begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} s_{11} \\ s_{21} \\ s_{31} \end{pmatrix} \quad \text{i analogicznie dla punktów } 010 \text{ i } 001$$

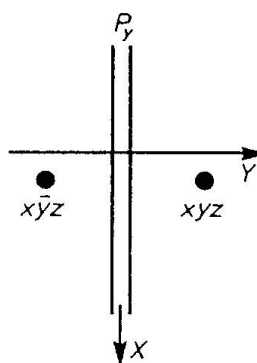
Rozważyć należy symetrie względem punktu, prostej i płaszczyzny. Każdej symetrii odpowiada element symetrii – zbiór punktów niezmienniczych danego przekształcenia. Symetria względem punktu to symetria środka **C**. Obrazem punktu w symetrii środka jest punkt równo odległy od środka przekształcenia, leżący na wspólnej prostej (**Rys. 1.8**). Zbiór punktów niezmienniczych jest jednoelementowy i zawiera tylko środek symetrii. Symetria środka wiąże 2 punkty równoważne symetrycznie.



**Rys. 1.8.**



**Rys 1.9**



**Rys. 1.10**

Jeśli założymy, że początek układu pokrywa się z położeniem środka symetrii, to symetrię środka opisać można równaniem:



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix}$$

Symetria względem prostej (**Rys. 1.9**) odpowiada obrotowi wokół osi dwukrotnej  $L_2$  o kąt  $180^\circ$ . Symetria osi dwukrotnej wiąże 2 punkty równoważne symetrycznie. Zbiór punktów niezmienniczych odpowiada prostej - osi obrotu. Ponieważ oś obrotu jest elementem kierunkowym, opisuje się go symbolem  $[uvw]$ . Macierzowe operatory dla podstawowych osi dwukrotnych mają postać:

$$L_2[100] \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad L_2[010] \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad L_2[001] \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Symetria względem płaszczyzny  $P$  pokazana jest na **Rys. 1.10**. Tak jak dla osi obrotu, położenie płaszczyzny opisuje się symbolem Millera  $(hkl)$ . Symetria płaszczyzny wiąże 2 punkty równoważne symetrycznie. Opis przekształceń przy pomocy operatorów macierzowych przedstawiony jest poniżej:

$$P(100) \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P(010) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P(001) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Dodatkowo występuje przekształcenie tożsamościowe  $E$  opisane operatorem:

$$E \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Definiuje się obroty o kąt  $360^\circ/n$  odpowiadające obrotowi wokół osi  $n$ -krotnej. Można wykazać, że w periodycznej sieci krystalicznej dozwolone są obroty o  $360/0$  180, 120, 90 i 60 stopni, odpowiadające osiom obrotu jednokrotnej  $L_1$ , dwukrotnej  $L_2$ , trzykrotnej  $L_3$ , czterokrotnej  $L_4$  i sześciokrotnej  $L_6$ . Jak widać, dolny indeks oznacza krotność osi obrotu. Symetria osi  $n$ -krotnej wiąże  $n$  punktów równoważne symetrycznie.

Przekształcenia złożone powstające przez połączone działanie obrotu wokół osi i symetrii środka (inwersji) noszą nazwę osi  $n$ -krotnych inwersyjnych  $L_{ni}$ . W zależności od krotności osi, liczba punktów równoważnych symetrycznie wynosi 2, 2, 6, 4 i 6 dla osi inwersyjnych jedno-, dwu-, trzy- cztero- i sześciokrotnych.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
 realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Rozkład materii w skończonej, ograniczonej przestrzeni (cząsteczka, komórka elementarna) można opisać przy pomocy grup punktowych symetrii, których elementami są symetrie opisane powyżej. Zbiory symetrii wraz z działaniem grupowym  $*$  tworzą grupy przemienne, jeśli spełniają warunki:

- Złożenie dwóch symetrii daje symetrię należącą do grupy
- Łączność. Dla każdego  $a, b, c \in G$   $(a*b) * c = a * (b*c)$
- Element neutralny  $e \in G$  taki, że dla każdego  $a \in G$   $a*e = e*a = a$
- El. przeciwny. Dla każdego  $a \in G$  istnieje  $a^{-1} \in G$  taki, że  $a*a^{-1} = a^{-1}*a = e$
- Przemienność. Dla każdego  $a, b \in G$  zachodzi  $a*b = b*a$

Złożenie  $S1 * S2 = S3$  odpowiada iloczynowy operatorów symetrii. Zamiast rachunku macierzowego można posłużyć się regułami składania operacji symetrii opisanymi poniżej:

Reguła 1 opisuje złożenie parzystokrotnych osi obrotu, prostopadłej płaszczyzny i środka symetrii. Obecność dwóch spośród trzech symetrii implikuje istnienie trzeciej.

$$L_{2n} * C = P \perp L_{2n} \quad P * C = L_2 \perp P \quad L_{2n} * P \perp = C$$

Reguła 2 dotyczy złożenia działania osi  $n$ -krotnej  $s$  działaniem osi dwukrotnych prostopadłych do  $L_n$ . Obecność jednej osi dwukrotnej implikuje istnienie  $n$  osi  $L_2$  prostopadłych do  $L_n$ . Kąt między sąsiadującymi osiami wynosi  $180^\circ/n$ . Istnienie osi  $L_2$  przecinających się pod kątem  $180^\circ/n$  implikuje istnienie osi  $L_n$  prostopadłej do płaszczyzny zawierającej osie dwukrotne.

$$L_n * L_2 \perp \Rightarrow nL_2 \perp L_n$$

Reguła 3 wskazuje, że obecność płaszczyzny symetrii równoległej do osi  $L_n$  oznacza istnienie  $n$  takich płaszczyzn, przecinających się pod kątem  $180^\circ/n$ . Także istnienie płaszczyzn przecinających się pod kątem  $180^\circ/n$  oznacza obecność osi  $L_n$  na linii ich przecięcia.

$$L_n * P \parallel \Rightarrow nP \parallel L_n$$

Reguła 4 wskazuje, że grupę tworzy parzystokrotna oś inwersyjna  $L_{2ni}$  oraz  $n$  osi  $L_2 \perp L_{2n}$  i  $n$  płaszczyzn równoległych do osi inwersyjnej. Liczba osi  $L_2$  i płaszczyzn jest równa połowie krotności osi inwersyjnej, położone są naprzemiennie, a kąt między nimi wynosi  $180^\circ/n$ .

$$L_{2ni} * L_2 \perp \text{ lub } L_{2ni} * P \parallel \Rightarrow nL_2 \perp L_{2ni} + nP \parallel$$

Zbiory elementów symetrii tworzące grupy przemienne nazywamy grupami punktowymi (klasami). Grupy punktowe opisują symetrię skończonej przestrzeni - cząsteczek lub komórki elementarnej. W przestrzeni  $E^3$  występują 32 grupy punktowe

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Do opisu symetrii stosuje się trzy symboliki. Pierwsza to symbolika Grota, podająca spis wszystkich symetrii działających w przestrzeni kryształu (osi obrotu  $L_n$ , inwersyjnych  $L_{ni}$ , płaszczyzn symetrii  $P$ , środka symetrii  $C$  i tożsamości  $E$ ).

Druga, symbolika Schoenflisa opiera się na założeniu, że główna oś obrotu jest pionowa (wertykalna). Jeśli występuje tylko jedna oś obrotu  $L_n$ , to symbol grupy to  $C_n$ . Jeśli zgodnie z regułą 2 występuje oś  $L_n$  i  $n L_2 \perp$  to taka grupa ma symbol  $D_n$ . Położenie płaszczyzn opisuje się literą w indeksie dolnym. Płaszczyzna prostopadła do osi  $n$ -krotnej (horyzontalna) daje symbol  $C_{nh}$ . Płaszczyzny równoległe do osi  $L_n$  (wertykalne, Reguła 3) dają symbol grupy  $C_{nv}$ .  $D_{nh}$  oznacza obecność płaszczyzny prostopadłej do głównej osi, co może oznaczać także płaszczyzny prostopadłe do osi  $L_2$ , czyli równoległe do  $L_n$ , jeśli jest to oś parzystokrotna (Reguła 1). Są też symbole osi inwersyjnych  $C_{ni}$ . Symbole specjalne dla grupy ze środkiem symetrii i lub z płaszczyzną to  $C_i$  i  $C_s$ . Symbole dla grup punktowych układu regularnego to  $T$ ,  $T_h$ ,  $T_d$ ,  $O$ ,  $O_h$ .

W krytalografii najwygodniejsza jest symbolika międzynarodowa, która daje spójne zasady opisu grup punktowych, jak i ich uogólnienia – grup przestrzennych. W symbolice międzynarodowej główne kierunki symetrii są opisane na pierwszym miejscu symbolu. Kierunki równoważne są opisane na tym samym miejscu symbolu. Elementy symetrii charakteryzujące ten sam kierunek występują na tym samym miejscu w symbolu, np.  $2/m$  oznacza symetrię osi dwukrotnej i prostopadłej do niej płaszczyzny. Zasady symboliki zależą od układu krytalograficznego:

1. Układ trójskośny - brak elementów kierunkowych:  $1$ ;  $\bar{1}$
2. Jednoskośny:  $\alpha = \gamma = 90^\circ < \beta$ :  $L_2 \parallel y$ ;  $P \perp y$   
W symbolu symetria kierunku  $[010]$ :  $2$ ,  $m$ ,  $2/m$
3. Rombowy -  $\alpha = \beta = \gamma = 90^\circ$ :  $L_2 \parallel x, y, z$ ;  $P \perp x, y, z$   
W symbolu symetria kierunku  $[100]$ ,  $[010]$ ,  $[001]$ :  
 $222$ ,  $mm2$ ,  $2/m2/m2/m$
4. Tetragonalny:  $a = b \neq c$ ;  $\alpha = \beta = \gamma = 90^\circ$ :  $L_4 \parallel z$ ;  $L_2 \parallel x, y$ ;  $P \perp x, y, z$   
W symbolu symetria kierunku  $z$ ,  $xy$ ,  $[110]$  i  $[\bar{1}10]$ :  
 $4$ ;  $4/m$ ;  $422$ ;  $4mm$ ;  $4/mmm$ ;  $\bar{4}$ ;  $\bar{4}2m$
5. Regularny:  $a = b = c$ :  $L_4, L_2 \parallel x, y, z$ ;  $L_3 \parallel [111]$ ;  $P \perp x, y, z; [110]$   
W symbolu symetria kierunków  $xyz$ ,  $[111]$ ,  $[110]$ :  
 $23$ ;  $2/m3$ ;  $432$ ;  $4/m32/m$ ;  $\bar{4}3m$
6. Trygonalny:  $a = b = c$ ;  $\alpha = \beta = \gamma \neq 90^\circ$ :  $L_3 \parallel z$ ;  $L_2 \parallel x_1, x_2, x_3$ ;  $P \perp x_i$   
W symbolu symetria kierunków  $z$ ,  $x_i$ , przekątnych:  
 $3$ ;  $32$ ;  $3m$ ;  $\bar{3}$ ;  $\bar{3}m$
7. Heksagonalny:  $a = b \neq c$   $\gamma = 120^\circ$ :  $L_6 \parallel z$ ;  $L_2 \parallel x_1, x_2, x_3$ ;  $P \perp x_i, z$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

W symbolu symetria kierunków  $z$ ,  $x_i$ , przekątnych:

$6$ ;  $6/m$ ;  $622$ ;  $6mm$ ;  $6/m2/m2/m$ ;  $\bar{6}$ ;  $\bar{6}2m$

Rzeczywiste kryształy są obiektami o budowie periodycznej, z translacyjnie powielaną komórką elementarną. Dlatego opis rozkładu materii w kryształach wymaga zastosowania translacyjnych elementów symetrii, które są uogólnieniem elementów symetrii opisanych powyżej. Translacyjne elementy symetrii odpowiadają przekształceniom złożonym z symetrii i translacji o wektor ułamkowy równoległy do elementu symetrii. Wyróżnia się dwa rodzaje translacyjnych elementów symetrii – osie śrubowe i płaszczyzny z poślizgiem.

Osie śrubowe o symbolu  $n_m$  definiuje się jako złożenie obrotu o kąt  $360^\circ/n$  i translacji o wektor  $m/n$  jednostkowej translacji wzdłuż osi obrotu. Przykładowo symbol osi  $6_5[001]$  oznacza obrót o  $60^\circ$  połączony z translacją o  $5/6 c_0$ . Należy zauważyć, że  $m$  jest liczbą naturalną o wartości od 1 do  $n-1$ . Wartości  $m=0$  i  $m=n$  odpowiadają zerowemu lub jednostkowemu przesunięciu czyli zwykłej osi obrotu.

Płaszczyzny z poślizgiem opisuje się symbolami literowymi. Płaszczyzny typu  $a$ ,  $b$  i  $c$  to połączenie symetrii płaszczyznowej z translacjami o wektor  $a/2$ ,  $b/2$  i  $c/2$  równoległy do płaszczyzny symetrii. Płaszczyzny o symbolu  $n$  to połączenie symetrii płaszczyznowej z translacjami o wektor  $(a+b)/2$ ,  $(b+c)/2$  i  $(a+c)/2$ . Płaszczyzny  $d$  to połączenie symetrii płaszczyznowej z translacjami o wektor  $(a+b)/4$ ,  $(b+c)/4$  i  $(a+c)/4$ . Wyjątkowo występują płaszczyzny  $n$  i  $d$  o translacjach  $(a+b+c)/2$  i  $(a+b+c)/4$ .

Operator symetrii opisujący te uogólnione przekształcenia ma postać macierzy  $4 \times 4$ , gdzie blok  $3 \times 3$  odpowiada symetrii, zaś  $t_x$ ,  $t_y$  i  $t_z$  to składowe translacje:

$$\begin{pmatrix} s_{11} & s_{12} & s_{13} & t_x \\ s_{21} & s_{22} & s_{23} & t_y \\ s_{31} & s_{32} & s_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

W roku 1891 niezależnie od siebie, Rosjanin Jewgraf Stiepanowicz Fiodorow i Niemiec Arthur Moritz Schönflies wyprowadzili wszystkie możliwe grupy przestrzennej symetrii, których w trójwymiarowej przestrzeni jest 230. Elementami grup przestrzennych są zwykła symetria, symetrie translacyjne oraz translacje opisane w grupach translacyjnych Bravais.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Wykład 2: Dyfrakcja promieniowania rentgenowskiego na kryształach

W warunkach laboratoryjnych można generować promieniowanie rentgenowskie stosując lampy rentgenowskie. W polu elektrycznym o regulowanej różnicy potencjałów, elektrony są przyspieszane między anodą i katodą lampy. Przy odpowiedniej różnicy potencjałów pola, elektrony trafiając w atomy katody mogą powodować wybicie elektronów z wewnętrznych powłok elektronowych oraz emisję promieniowania. Zjawisko można opisać w oparciu o zasadę zachowania energii uwzględniając energię kinetyczną elektronu przyspieszanego  $E_{el}$  i energię kinetyczną elektronów po zderzeniu  $E_{el} = E_1 + E_2 + h\nu$ . Promieniowanie generowane podczas zderzenia daje widmo ciągłe. Po wybiciu elektronów z wewnętrznych powłok następuje spontaniczne przeniesienie elektronów z powłok wyższych, połączone z emisją promieniowania o energii  $h\nu = \Delta E$  odpowiadającej różnicy energii poziomów, a więc charakterystycznej dla materiału anody. To promieniowanie określa się mianem widma charakterystycznego. Uzyskuje się promieniowanie rentgenowskie o użytecznej w krystalografii długości fali  $\lambda$  0.5 - 2.5 Å. To promieniowanie rentgenowskie oddziałuje z materią przez wzbudzenie atomów związane z przeniesieniem elektronów na wyższe poziomy energetyczne. Przez analogię do optyki, zakres długości fali  $\lambda$  0.5 - 2.5 Å pozwala na ‘zobaczenie’ szczegółów rozkładu atomów z rozdzielczością definiowaną jako  $\lambda/2$ . Przykładowo, zastosowanie lampy z anodą Cu daje promieniowanie  $\lambda=1.54178$  Å, a anoda Mo generuje promieniowanie  $\lambda=0.79021$  Å. Porównanie z długościami wiązań C-C 1.54 Å, C=O 1.28 Å, C-H 1.08 Å, O-H 0.92 Å wskazuje, że użycie promieniowania rentgenowskiego pozwala zobaczyć cząsteczki chemiczne jako zbiór atomów.

Oddziaływanie promieniowania rentgenowskiego z atomami jest związane z liczbą elektronów danego atomu. Ponieważ efekt jest addytywny, wkład każdego atomu danego pierwiastka jest jednakowy i wynika z liczby atomowej tego pierwiastka (**Rys. 2.1**). Opisuje się go przez atomowy czynnik rozpraszania  $f_i$ , wyrażony równaniem

$$f_i = f_0 \exp(-K \sin^2 \theta / \lambda^2), \text{ gdzie}$$

$f_0$  atomowy czynnik rozpraszania atomu nie wykonującego oscylacji

$K$  czynnik temperaturowy zależny od amplitudy oscylacji atomu wokół położenia równowagi

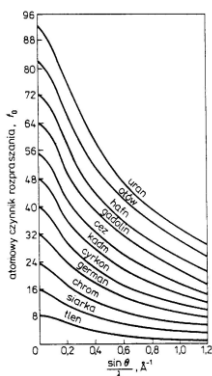
$\lambda$  długość fali promieniowania

$\theta$  kąt pod jakim obserwuje się promieniowanie ugięte na atomie

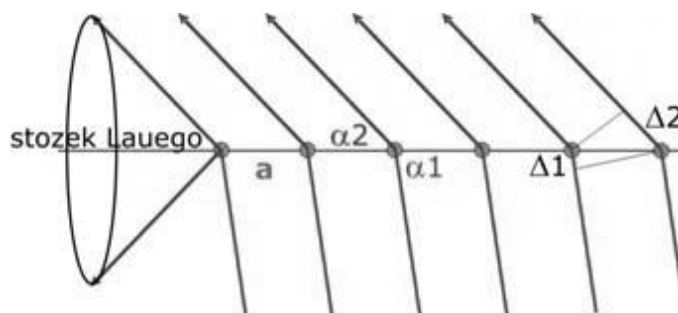
Sieć krystaliczna wykazuje periodyczną budowę. Zbiory cząsteczek (materia) definiują proste i płaszczyzny w sieci, których opis można podać w odniesieniu do wybranego układu osi, czyli układu krystalograficznego. Stosowana w krystalografii symbolika [uvw] i (hkl) odnosi się do wszystkich prostych lub płaszczyzn równoległych w przestrzeni.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Opis dyfrakcji promieniowania na węzłach (cząsteczkach) podał Laue. W przypadku jednowymiarowej sieci przestrzennej, na prostą o stałą  $a$  kierujemy wiązkę pod kątem padania  $\alpha_1$ . Wiązki oddziałują z elektronami kolejnych cząsteczek, stając się źródłem fal kulistych (**Rys. 2.2**). Fale interferują dając wzmocnienie przy różnicy dróg optycznych  $\Delta_2 - \Delta_1$  równej całkowitej wielokrotności długości fali, co opisuje równanie Lauego  $h\lambda = a(\cos\alpha_2 - \cos\alpha_1)$ . W równaniu współczynnik  $h$  oznacza rząd wzmocnienia wiązki.



Rys. 2.1



Rys. 2.2

Równanie Lauego wskazuje na ścisły związek geometrii sieci (stała  $a$ ) i geometrii dyfrakcji (kąty  $\alpha_2$  i  $\alpha_1$ ). Dla sieci trójwymiarowej refleks wymaga spełnienia układu trzech równań:

$$h\lambda = a(\cos\alpha_2 - \cos\alpha_1) \quad k\lambda = b(\cos\beta_2 - \cos\beta_1) \quad l\lambda = c(\cos\gamma_2 - \cos\gamma_1)$$

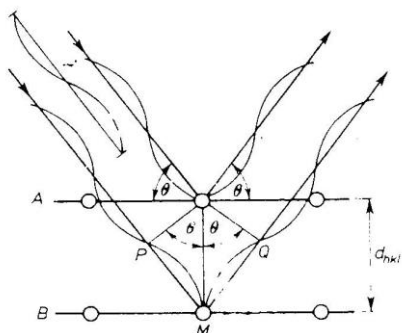
Alternatywny (i równoważny) opis dyfrakcji z wykorzystaniem prostego formalizmu odbiciowego został wprowadzony przez Braggów. Wiązka promieniowania pada na rodzinę płaszczyzn ( $hkl$ ), które działają jak zwierciadło odbijając promieniowanie, a kąt padania  $\theta$  jest równy kątowi odbicia (**Rys. 2.3**). Przy zgodności fazy wiązek odbitych następuje interferencja ze wzmocnieniem wiązki, zaś przy różnicy faz wiązek interferujących następuje osłabienie, które osiąga maksymalną wartość dla różnicy faz  $180^\circ$ . Maksymalne wzmocnienie następuje, gdy różnica dróg optycznych  $\Delta = PM + MQ$  równa jest całkowitej wielokrotności długości fali. Równanie Bragga formułuje warunek na wzmocnienie wiązki promieniowania o długości fali  $\lambda$  padającego pod kątem  $\theta$  na płaszczyznę ( $hkl$ ) o odległości  $d_{hkl}$ :

$$n\lambda = 2d_{hkl} \sin\theta \quad \text{lub} \quad d_{hkl}/n = 2\sin\theta / \lambda$$

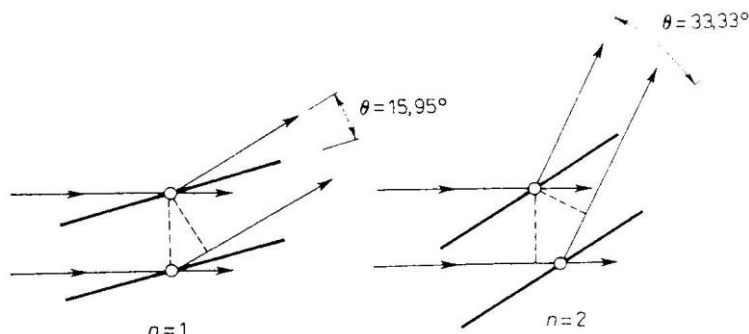
Współczynnik  $n$  to rząd odbicia. Z równania Bragga wynika, że wzmocnienie następuje pod kątami zależnymi od odległości międzypłaszczyznowej  $d_{hkl}$ . Oznacza to, że kąty Bragga  $\theta$ , pod którymi obserwujemy wzmocnienie zależą ściśle od geometrii sieci krystalicznej. Wiązkę wzmocnioną po odbiciu od płaszczyzn ( $hkl$ ) oznaczamy symbolem  $hkl$ . Konsekwencją opisu jest też, że każda rodzina płaszczyzn ( $hkl$ ) daje wzmocnienia odpowiadające kolejnym wielokrotnościom długości fali promieniowania (**Rys. 2.4**). Wiązkę drugiego rzędu oznaczmy symbolem  $2h, 2k, 2l$  itd.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Dla płaszczyzn (231) obserwować będziemy serię wzmocnień 1, 2, 3 i kolejnych rzędów o symbolach 231, 462, 693 itd. W opisie można uniezależnić się od rzędu odbicia dzieląc  $d' = d_{hkl}/n = 2\sin\theta / \lambda$ .



Rys. 2.3



Rys. 2.4

Sieć krystaliczna obok periodyczności charakteryzuje się występowaniem symetrii. Opiszmy zjawisko dyfrakcji dla dwóch rodzin płaszczyzn związanych symetrią. Załóżmy, że rozważamy występowanie symetrii osi  $L_2[010]$ . Wtedy płaszczyzny  $(hkl)$  i  $(-hk-l)$  są równoważne, co oznacza równość odległości  $d_{hkl}$  i  $d_{-hk-l}$ . Z równania Bragga wynika, że kąty Bragga dla obu rodzin płaszczyzn są identyczne:

$$2\sin\theta / \lambda = d'_{hkl} = d'_{-hk-l} = 2\sin\theta / \lambda$$

Obraz dyfrakcyjny wykaże symetrię związaną z symetrią sieci krystalicznej. Natężenie wiązki ugiętej od rodziny płaszczyzn  $(hkl)$  kryształu  $I_{hkl}$  związane jest z czynnikami struktury  $F_{hkl}$ , które można wyliczyć jako sumę wkładów atomów tworzących strukturę:

$$I_{hkl} \sim F_{hkl} = \sum_{i=1}^N f_i \exp 2\pi i(hx_i + ky_i + lz_i)$$

$F$  czynnik struktury,  $\sim \sqrt{I_{hkl}}$

$f_i$  atomowy czynnik rozpraszania

$h, k, l$  wskaźniki płaszczyzny dającej wiązkę  $I_{hkl}$

$x, y, z$  współrzędne  $i$ -tego atomu w komórce elementarnej

Natężenie  $I_{hkl}$  wiązki ugiętej zależy od rodzaju atomu ( $f_i$ ) i położenia ( $xyz$ ), Geometria dyfrakcji zależy od parametrów sieci ( $d_{hkl}, a, b, c, \alpha, \beta, \gamma$ ), a te są powiązane z rozmiarami atomów i cząsteczek tworzących sieć krystaliczną oraz oddziaływań między nimi.

Można obliczyć wartości czynników struktury dla poprzedniego przykładu kryształu o symetrii  $L_2[010]$ . Atomy równoważne symetrycznie mają współrzędne  $xyz$  i  $-xy-z$ . Czynniki struktury pochodzące od płaszczyzn związanych symetrią  $F_{hkl}$  i  $F_{-hk-l}$  wyrażone jako suma wkładów pary atomów równoważnych symetrycznie mają równe wartości:

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
 realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

$$F_{hkl} = f_i [\exp 2\pi i \{hx + ky + lz\} + \exp 2\pi i \{h(-x) + ky + l(-z)\}]$$

$$F_{-h-k-l} = f_i [\exp 2\pi i \{(-h)x + ky + (-l)z\} + \exp 2\pi i \{(-h)(-x) + ky + (-l)(-z)\}] = F_{hkl}$$

Zarówno geometria rozkładu wiązek jak i natężenia wykazują symetrię związaną z symetrią rozkładu materii w kryształach - symetrią sieci. Jednak formalizm opisu dyfrakcjami wiąże się z prawem Friedla. Rozważmy czynniki struktury związane z refleksami ugiętymi na płaszczyznach (hkl) i (-h-k-l), nie robiąc żadnych założeń o symetrii kryształu. Wykorzystamy twierdzenie Eulera dla funkcji eksponencjalnej z wykładnikiem urojonym mówiące, że  $\exp(i\alpha) = \cos\alpha + i \sin\alpha$  oraz parzystość funkcji trygonometrycznych.

$$F_{hkl} = \sum f_i \exp 2\pi i (hx+ky+lz) = \sum f_i [\cos 2\pi (hx+ky+lz) + i \sin 2\pi (hx+ky+lz)]$$

$$F_{hkl} = \sum f_i \cos 2\pi (hx+ky+lz) + i \sum f_i \sin 2\pi (hx+ky+lz) = A_{hkl} + i B_{hkl}$$

$$I_{hkl} \sim F_{hkl} F_{hkl}^* = (A + iB) (A - iB) = A_{hkl}^2 + B_{hkl}^2$$

$$F_{-h-k-l} = \sum f_i \exp 2\pi i (-hx-ky-lz) = \sum f_i [\cos 2\pi (-hx-ky-lz) + i \sin 2\pi (-hx-ky-lz)]$$

$$F_{-h-k-l} = \sum f_i \cos 2\pi (hx+ky+lz) + i \sum f_i \sin 2\pi (hx+ky+lz) = A_{hkl} - i B_{hkl}$$

$$I_{-h-k-l} \sim F_{-h-k-l} F_{-h-k-l}^* = (A_{hkl} - i B_{hkl}) (A_{hkl} + i B_{hkl}) = A_{hkl}^2 + B_{hkl}^2$$

Stąd  $I_{hkl} = I_{-h-k-l}$

Z prawa Friedla wynika, że natężenia refleksów hkl i -h-k-l są jednakowe, czyli obraz dyfrakcyjny jest centrosymetryczny (wykazuje symetrię środka), bez względu na rzeczywistą symetrię sieci krystalicznej. W konsekwencji symetria obrazu dyfrakcyjnego odpowiada symetrii 32 grup punktowych ale zawiera symetrię środka, co ogranicza liczbę możliwości do 11 grup dyfrakcyjnych Lauego:

Grupa punktowa	Grupa dyfrakcyjna Lauego
1, $\bar{1}$	$\bar{1}$
2, m, 2/m	2/m
222, mm2, mmm	mmm
3, $\bar{3}$	$\bar{3}$
32, 3m, $\bar{3}m$	$\bar{3}m$
4, $\bar{4}$ , 4/m	4/m
422, 4mm, 4/mmm, $\bar{4}2m$	4/mmm
6, $\bar{6}$ , 6/m	6/m
622, 6mm, 6/mmm, $\bar{6}2m$	6/mmm
23, m3	m3
432, $\bar{4}3m$ , m3m	m3m



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Istotną cechą jest, że obraz dyfrakcyjny niesie dodatkową informację. Tzw. wygaszenia systematyczne oznaczają zerowe natężenie refleksów i czynników struktury  $F_{hkl}$  w klasach refleksów, niezależne od rodzaju atomów i ich położenia w komórce elementarnej. Wartość czynnika struktury oblicza się jako sumę wkładów atomów związanych przekształceniem, stosując wspomniane już twierdzenie Eulera.

Rozważmy wpływ sieci translacyjnych Bravais na natężenia wiązek. W komórce prymitywnej atom o współrzędnych  $xyz$  ma atom równoważny w sąsiednich komórkach (przesunięcia jednostkowe lub ich kombinacja liniowa). Łatwo wykazać, że czynnik struktury  $F_{hkl} = f_i \exp 2\pi i(hx+ky+lz)$  nie przyjmuje wartości zero niezależnie od położenia atomów.

Wprowadźmy ułamkową translację  $(a+b)/2$  definiującą sieć translacyjną C. Współrzędne atomów równoważnych to  $xyz$  oraz  $1/2+x, 1/2+y, z$ . Dla dowolnych refleksów  $hkl$ :

$$F_{hkl} = f_i [\exp 2\pi i(hx+ky+lz) + \exp 2\pi i(h(1/2+x)+k(1/2+y)+lz)]$$

$$F_{hkl} = f_i [\exp 2\pi i(hx+ky+lz) + \exp 2\pi i(hx+ky+lz) \exp 2\pi i(h/2 + k/2)]$$

$$F_{hkl} = f_i \exp 2\pi i(hx+ky+lz) [1 + \exp 2\pi i(h/2 + k/2)]$$

$$F_{hkl} = f_i \exp 2\pi i(hx+ky+lz) [1 + \exp \pi i(h+k)]$$

$$F_{hkl} = f_i \exp 2\pi i(hx+ky+lz) [1 + \cos \pi (h+k) + i \sin \pi (h+k)]$$

Rozważmy wartość czynnika w nawiasie kwadratowym. Dla wskaźników  $h+k \in C$  wartość  $\sin \pi (h+k) = 0$ . Stąd wyrażenie upraszcza się dając warunek na wygaszenia systematyczne:

- dla  $h+k=2n+1$      $1+\cos \pi (h+k)=0$      $F_{hkl} = 0$
- dla  $h+k=2n$      $1+\cos \pi (h+k)=2$      $F_{hkl} = 2f_i \exp 2\pi i(hx+ky+lz) = 2 F_{hkl} (P)$

Obserwowane czynniki struktury dla sieci C mają wartość dwukrotnie większą niż hipotetyczne czynniki od sieci prymitywnej. Jednocześnie czynniki struktury dla nieparzystej sumy  $h+k$  mają wartość zero. Wyliczony efekt związany jest z ułamkową translacją występującą w sieci krystalicznej. Można przeprowadzić analogiczne wyliczenia dla innych typów sieci translacyjnych i związanych z nimi wygaszeń systematycznych:

Typ sieci translacyjnej Bravais	Wygaszenia systematyczne
P	brak
A	$k+l=2n+1$
B	$h+l=2n+1$
C	$h+k=2n+1$
F	$h+k=2n+1, h+l=2n+1, k+l=2n+1$
I	$h+k+l=2n+1$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
 realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

W 230 grupach przestrzennych występują symetrie zarówno zwykłe jak i translacyjne. Można wykazać, że zwykła symetrie nie dają wygaszeń systematycznych ze względu na brak ułamkowych translacji z nimi związanych. Jednak osie śrubowe i płaszczyzny z poślizgiem dają wygaszenia w klasach refleksów, pozwalając na identyfikację zarówno orientacji jak i rodzaju elementy symetrii. Poniższy przykład ilustruje to zagadnienie. Dla osi  $2_1||y$  współrzędne punktów równoważnych symetrycznie to  $xyz$  oraz  $-x, 1/2+y, -z$ . Suma wkładów obu atomów:

$$F_{hkl} = f_i [\exp 2\pi i(hx+ky+lz) + \exp 2\pi i(h(-x)+k(1/2+y)+l(-z))]$$

$$F_{hkl} = f_i \exp 2\pi i(hx+ky+lz)[1 + \exp 2\pi i(-2hx+k/2-2lz)]$$

Dla dowolnych  $xyz$  trzeba przyjąć  $h,l=0$  i rozważać refleksy  $0k0$ :

$$F_{hkl} = f_i \exp 2\pi i(hx+ky+lz)[1 + \cos k\pi]$$

Warunek na wygaszenia systematyczne w klasie refleksów  $0k0$ :  $F_{hkl} = 0$  gdy  $\cos k\pi = -1$  czyli dla  $k=2n+1$ . W przeciwnym wypadku dla  $k=2n$  wartość  $\cos k\pi = 1$  a  $F_{hkl} \neq 0$ . Klasa refleksów z wygaszeniami zależy od orientacji elementów symetrii, zaś warunek na wygaszenia związany jest z wektorem ułamkowej translacji. Podsumowuje to tabela:

Orientacja	Element symetrii	Wygaszenia systematyczne
dowolna	$m, 2, 3, 4, 6$	brak
[100]	$2_1$	$h00; h=2n+1$
[010]	$2_1$	$0k0; k=2n+1$
[001]	$2_1$	$00l; l=2n+1$
(100)	$b$	$0kl; k=2n+1$
	$c$	$0kl; l=2n+1$
	$n$	$0kl; k+l=2n+1$
(010)	$a$	$h0l; h=2n+1$
	$c$	$h0l; l=2n+1$
	$n$	$h0l; h+l=2n+1$
(001)	$a$	$hk0; h=2n+1$
	$b$	$hk0; k=2n+1$
	$n$	$hk0; h+k=2n+1$

Analiza wygaszeń systematycznych jest więc metodą określenia rzeczywistej grupy przestrzennej kryształu. Pozwala ona na rozwiązania niejednoznaczności wynikającej z prawa Friedla gdy wygaszenia systematyczne wykażą obecność parzystokrotnej osi śrubowej i prostopadłej do niej płaszczyzny z poślizgiem.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

### Wykład 3: Teoria czynników struktury i problem fazowy

Trójwymiarową periodyczną sieć krystaliczną opisujemy w bazie wektorów  $\mathbf{a}$ ,  $\mathbf{b}$  i  $\mathbf{c}$ , związanych z wybranym układem krystalograficznym. Jednak w przestrzeni można zdefiniować nieskończenie wiele baz i dokonywać transformacji między nimi. W szczególności dla bazy wektorów  $\mathbf{a}$ ,  $\mathbf{b}$  i  $\mathbf{c}$  przestrzeni rzeczywistej możemy wybrać bazę wektorów odwrotnych  $\mathbf{a}^*$ ,  $\mathbf{b}^*$  i  $\mathbf{c}^*$  definiujących sieć odwrotną. Wektory sieci odwrotnej można wybrać następująco: dla płaszczyzn rodzin (100), (010) i (001) prowadzi się proste prostopadłe, z punktu wybranego za początek układu w sieci odwrotnej odmierzając odcinek  $1/d_{100}$  równy  $\mathbf{a}^*$  i odpowiednie dla pozostałych osi. Obrazem płaszczyzny (hkl) w przestrzeni rzeczywistej jest punkt hkl w przestrzeni odwrotnej definiowany przez wektor  $\mathbf{h} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$ .

Funkcję  $G(\mathbf{h})$  nazywamy transformatą Fouriera funkcji  $g(\mathbf{x})$ , gdy

$$G(\mathbf{h}) = T[g(\mathbf{x})] = \int_{V_{\mathbf{x}}} g(\mathbf{x}) \cdot \exp(2\pi i \mathbf{h} \cdot \mathbf{x}) dV_{\mathbf{x}}$$

Wektor  $\mathbf{x}$  definiowany jest w przestrzeni rzeczywistej wektorów  $\mathbf{a}$ ,  $\mathbf{b}$  i  $\mathbf{c}$ , zaś  $\mathbf{h}$  w przestrzeni odwrotnej  $\mathbf{a}^*$ ,  $\mathbf{b}^*$  i  $\mathbf{c}^*$ . Funkcja  $g(\mathbf{x})$  jest wyznaczana z odwrotnej transformaty Fouriera funkcji  $G(\mathbf{h})$ :

$$g(\mathbf{x}) = T^{-1}[G(\mathbf{h})] = k \int_{V_{\mathbf{h}^*}} G(\mathbf{h}) \cdot \exp(-2\pi i \mathbf{h} \cdot \mathbf{x}) dV_{\mathbf{h}^*}$$

Splotem dwóch funkcji  $g(\mathbf{x})$  i  $q(\mathbf{x})$  jest funkcja  $c(\mathbf{x})$  zdefiniowana jako:

$$c(\mathbf{x}) = g(\mathbf{x}) \otimes q(\mathbf{x}) = \int_{V_{\mathbf{x}}} g(\mathbf{x}') \cdot q(\mathbf{x} - \mathbf{x}') dV$$

Przesunięcie funkcji  $g(\mathbf{x})$  o wektor  $\mathbf{a}$  na osi odciętych nie zmienia jej części rzeczywistej (amplitudy), a jedynie część urojoną (fazę). Zmiana fazy jest proporcjonalna do  $\mathbf{h}$ :

$$T[g(\mathbf{x} - \mathbf{a})] = \exp(-2\pi i \mathbf{h} \cdot \mathbf{a}) \cdot G(\mathbf{h})$$

Jest kilka ważnych twierdzeń dotyczących transformat. Transformata sumy funkcji jest sumą ich transformat. Transformata splotu dwóch funkcji jest iloczynem ich transformat:

$$T[g(\mathbf{x}) \otimes q(\mathbf{x})] = G(\mathbf{h}) \cdot Q(\mathbf{h})$$

Transformata funkcji odwrotnej jest  $G^*(\mathbf{h}) = T[g^*(-\mathbf{x})]$ . Można pokazać, że

$$P(\mathbf{x}) = \int_{V_{\mathbf{h}^*}} G(\mathbf{h}) \cdot G^*(\mathbf{h}) \cdot \exp(-2\pi i \mathbf{h} \cdot \mathbf{x}) dV_{\mathbf{h}^*} = \int_{V_{\mathbf{h}^*}} |G(\mathbf{h})|^2 \cdot \exp(-2\pi i \mathbf{h} \cdot \mathbf{x}) dV_{\mathbf{h}^*}$$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

W komórce elementarnej można opisać rozkład gęstości elektronowej jako sumę rozkładów dla atomów  $a(\mathbf{x}) = \sum a_i(\mathbf{x}-\mathbf{x}_i)$ . Transformata sumy funkcji jest sumą ich transformat. Jeśli transformata  $T[a_i(\mathbf{x}-\mathbf{x}_i)] = f_i(\mathbf{h})$ , to transformatę funkcji  $a(\mathbf{x})$  można zapisać:

$$T[a(\mathbf{x})] = \sum f_i(\mathbf{h}) \cdot \exp(2\pi i \mathbf{h} \cdot \mathbf{x}_i)$$

W krystalografii funkcja  $g(\mathbf{x})$  to funkcja rozkładu gęstości elektronowej  $\rho(\mathbf{x})$ , zaś funkcja  $G(\mathbf{h})$  to czynnik struktury  $F(\mathbf{h})$ . Stąd można zapisać:

$$F(\mathbf{h}) = \int_{V_x} \rho(\mathbf{x}) \cdot \exp(2\pi i \mathbf{h} \cdot \mathbf{x}) dV_x$$

Dla dyskretnego rozkładu punktowych atomów:

$$F(\mathbf{h}) = \sum f_i \cdot \exp(2\pi i \mathbf{h} \cdot \mathbf{x}_i)$$

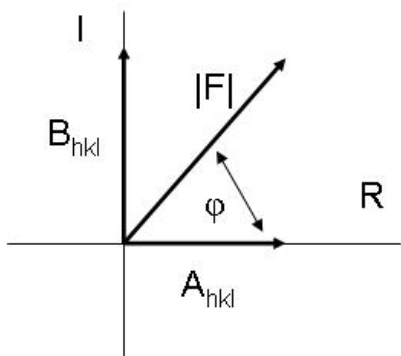
Odwrotna transformacja Fouriera pozwala wyliczyć funkcję rozkładu gęstości elektronowej  $\rho(\mathbf{x})$  w punkcie o współrzędnych  $\mathbf{x}$

$$\rho(\mathbf{x}) = V^{-1} \sum_{\mathbf{h}} F(\mathbf{h}) \cdot \exp(-2\pi i \mathbf{h} \cdot \mathbf{x})$$

Transformacja Fouriera z czynników struktury pozwala wyliczyć rozkładu gęstości elektronowej w kryształach  $\rho(xyz) = V^{-1} \sum F(hkl) \cdot \exp[-2\pi i(hx+ky+lz)]$ . Bezpośrednie użycie tego równania nie jest możliwe. Czynniki struktury, jako wielkości zespolone opisuje się amplitudą  $|F(hkl)|$  oraz fazą  $\varphi(hkl)$ :  $F(hkl) = |F(hkl)| \cdot \exp[i\varphi(hkl)]$ . Pomiar dyfrakcyjny dostarcza nam informacji o intensywnościach refleksów  $I(hkl)$ , które są proporcjonalne do kwadratu amplitudy  $|F(hkl)|^2$ . Można łatwo wyliczyć amplitudy jako pierwiastek z natężeń refleksów otrzymanych w eksperymencie dyfrakcyjnym  $|F(hkl)| = \sqrt{I(hkl)}$ , jednak brak informacji o fazie. To faza decyduje o jasności uzyskanej funkcji rozkładu gęstości elektronowej.

Problem fazowy można zilustrować **Rys. 3.1**. Czynniki struktury  $|F_{hkl}| \exp i\varphi$  można wyrazić jako sumę wektorową składowych rzeczywistej i urojonej.

$$F_{hkl} = |F_{hkl}| \exp i\varphi = |F_{hkl}|(\cos\varphi + i\sin\varphi) = A_{hkl} + iB_{hkl}$$



**Rys. 3.1**

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Nateżenie zmierzone w eksperymencie dyfrakcyjnym  $I_{hkl} \sim |F_{hkl}|^2 = FF^* = A^2 + B^2$ . Nieznaną fazę  $\varphi$  Można ją wyliczyć ze znanych położenia atomów:

$$F_{hkl} = |F_{hkl}|(\cos\varphi + i\sin\varphi) = A_{hkl} + iB_{hkl}$$

$$\operatorname{tg}\varphi_{hkl} = B_{hkl} / A_{hkl}$$

Wtedy eksperymentalne  $F_{hkl} = |F_{hkl}| \exp i\varphi$

lub w oparciu o model  $F = \sum f_i \exp 2\pi i(hx_i + ky_i + lz_i)$

Funkcja rozkładu gęstości elektronowej w punkcie o współrzędnych xyz

$$\rho_{xyz} = V^{-1} \sum_{hkl} F_{hkl} \exp -2\pi i(hx+ky+lz)$$

lub gdy znamy rodzaj i położenie atomów w komórce  $x_i, y_i, z_i$

$$\rho_{xyz} = V^{-1} \sum_{hkl} \left[ \sum_i f_i \exp 2\pi i(hx_i+ky_i+lz_i) \right] \exp -2\pi i(hx+ky+lz)$$

Istotą problemu fazowego jest więc fakt, że fazę niezbędną do określenia położenia atomów możemy wyliczyć znając te położenia.

Trzeba zauważyć, że w wielu przypadkach można odgadnąć położenie wybranych atomów w strukturze porównując liczbę punktów równoważnych symetrycznie w danej grupie przestrzennej (lprs) z liczbą cząsteczek w komórce elementarnej  $Z$  wyznaczoną w oparciu o pomiar gęstości kryształu  $d$ . Gęstość można wyrazić jako iloraz masy i objętości komórki elementarnej:

$$d = m_k / V_k = 1.6604 Z \text{ Mcz} / V_k$$

$$Z = 0.6023 d V_k / \text{Mcz}$$

Liczbę punktów równoważnych symetrycznie oblicza się dla każdej grupy przestrzennej jako iloczyn liczby punktów związanych w grupie translacyjnej Bravais i liczby punktów związanych symetriami w danej grupie. Należy przypomnieć, że liczba punktów równoważnych symetrycznie dla translacyjnych elementów symetrii jest równa liczbie punktów związanych nietranslacyjnymi symetriami.

Jeżeli  $Z > \text{lprs}$ , to w części asymetrycznej struktury mamy więcej niż 1 cząsteczkę (jednostkę chemiczną). Gdy  $Z = \text{lprs}$ , część asymetryczna struktury zawiera jedną cząsteczkę. W obu przypadkach cząsteczki mają tylko symetrię tożsamościową i nie można nic wywnioskować o położeniu cząsteczki. W przypadku gdy  $Z < \text{lprs}$ , ułamek cząsteczki równy ilorazowi  $Z/\text{lprs}$  znajduje

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

się w części asymetrycznej, a zatem cząsteczka wykazuje wewnętrzną symetrię. W grupach przestrzennych mamy symetrie translacyjne lub zwykłe, a tylko te ostatnie (osie obrotu, płaszczyzny, środek) mogą opisywać symetrię cząsteczki. Jeśli zatem w cząsteczce występuje atom, który musi być niezmienniczy w danym przekształceniu, to jego współrzędne są ściśle określone. Aby przybliżyć sposób rozumowania, rozważmy wybrane przykłady.

Kompleks  $[\text{Pt}(\text{NH}_3)_2\text{Cl}_2]$  krystalizuje w rombowej grupie przestrzennej  $Pbca$ . Zatem pełny zestaw symetrii obejmuje trzy płaszczyzny z poślizgiem, trzy osie śrubowe  $2_1$  oraz środek symetrii. Liczba punktów równoważnych symetrycznych w tej grupie  $l_{\text{pr}} = 8$ . Pomiar gęstości wykazał, że w objętości komórki elementarnej mamy 8 cząsteczek. Dlatego na część asymetryczną przypada  $8/8=1$  cząsteczka, co oznacza, że cząsteczka nie ma wewnętrznej symetrii. Jeśli pomiar gęstości wykaże, że  $Z=4$ , wtedy w części asymetrycznej mamy  $4/8$  czyli połowę cząsteczki. Drugą część cząsteczki uzyskamy przekształcając pierwszą stosując operację symetrii wiążącą 2 punkty równoważne symetrycznie. Wszystkie symetrie występujące w tej grupie przestrzennej (płaszczyzny z poślizgiem, osie śrubowe i środek symetrii) spełniają ten warunek. Jednak tylko symetria środka (nie translacyjna) może opisać symetrię cząsteczki badanego kompleksu. Ponieważ cząsteczka zawiera jeden jon centralny, musi on być punktem niezmienniczym przekształcenia, czyli jego położenie pokrywa się z położeniem środka. W konsekwencji znamy położenie jonu  $\text{Pt}(0,0,0)$

Kryształ  $[\text{Pt}(\text{NH}_3)_2\text{Cl}_2]$  otrzymano w grupie  $P2_1/c$  przy  $Z=2$ . W tej grupie mamy dwa translacyjne elementy symetrii – oś śrubową i płaszczyznę z poślizgiem, oraz nietranslacyjną symetrię środka. Tylko nietranslacyjne elementy symetrii mogą opisywać symetrię cząsteczki. Dlatego w omawianym przykładzie cząsteczka musi wykazywać symetrię środka, a pojedynczy jon platyny musi leżeć na środku symetrii. Przyjęto, że wyboru początku układu współrzędnych dokonuje się tak, że jest to jeden z punktów niezmienniczych przekształceń w grupie symetrii. Jeśli grupa jest centrosymetryczna, to początek układu pokrywa się ze środkiem symetrii. Stąd współrzędne jonu  $\text{Pt}^{2+}$  to 000.

Kompleks  $[\text{Pt}(\text{NH}_3)_2\text{Cl}_2]$  krystalizuje w grupie przestrzennej  $Pnn2$ . Jest to grupa z układu rombowego, a więc trzy miejsca w symbolu charakteryzują symetrię kierunków X, Y i Z układu. Oznacza to, że oś dwukrotna ma kierunek Z. W tej grupie są 4 położenia równoważne symetrycznie. Z pomiaru gęstości wynika wartość  $Z=2$ , co oznacza  $1/2$  cząsteczki w części asymetrycznej i w konsekwencji symetrię  $C_2$  cząsteczki. Oś dwukrotna to zbiór punktów  $\{0, 0, z\}$ , gdzie  $z$  przyjmuje dowolną wartość. Analizowana grupa należy do grup polarnych, i można przyjąć dowolną wartość współrzędnej  $z$  dla jednego z atomów, wyznaczając w ten sposób jego położenie względem początku układu, czyli wybierając początek układu! To pozwala wyliczyć wartość fazy  $\varphi$  i funkcję rozkładu gęstości elektronowej.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

W podobnych przypadkach wydedukowanie współrzędnych atomów pozwala na obliczenie przybliżonej wartości fazy i dalej funkcji rozkładu gęstości elektronowej. Takie odgadnięcie struktury jest jednym ze sposobów rozwiązania problemu fazowego.

Dla odgadniętego lub znanego położenia atomu lub kilku atomów, można wyliczyć czynnik struktury zgodnie ze wzorem:

$$F = \sum f_i \exp 2\pi i(hx_i + ky_i + lz_i) = A_{hkl} + iB_{hkl}$$

Fazy obliczone dla wszystkich refleksów są przybliżone, gdyż oparte są o wkład tylko kilku atomów. Jednak jeśli ten wkład jest istotny (dla ciężkich atomów  $f_i$  jest duże), to uzyskane wartości są dobrym przybliżeniem rzeczywistych faz. Wtedy maksima na obliczonej mapie rozkładu gęstości elektronowej ujawnią położenie kolejnych atomów, których wkład do wartości obliczonego czynnika struktury a w konsekwencji do fazy uwzględnia się w kolejnym obliczeniu. Tak metodą kolejnych iteracji można zlokalizować wszystkie ciężkie atomy, poprawiając wyliczone wartości fazy dla wszystkich refleksów. Znalezienie położenia atomów wodoru prowadzi się po uściśleniu struktury metodą najmniejszych kwadratów przez dopasowanie współrzędnych i amplitud oscylacji atomów niewodorowych do danych eksperymentalnych (obserwowane natężenia lub czynniki struktury)

## Wykład 4: Metody małowzrostkowe rozwiązywania problemu fazowego

Problem fazowy polega na tym, że z eksperymentu dyfrakcyjnego otrzymujemy wartości natężeń wiązek  $I_{hkl}$ , a te można przeliczyć na amplitudy czynników struktury  $|F_{hkl}|$ . Jednak te eksperymentalne dane nie pozwalają na obliczenie odwrotnej transformaty Fouriera

$$\rho_{xyz} = V^{-1} \sum_{hkl} F_{hkl} \exp -2\pi i(hx+ky+lz)$$

gdyż nie znamy wartości fazy pozwalającej wyliczyć  $F_{hkl} = |F_{hkl}| \exp i\varphi$ . Fazę można wyliczyć, jeśli znamy rodzaj i położenie atomów w komórce  $x_i, y_i, z_i$

$$\rho_{xyz} = V^{-1} \sum_{hkl} \left[ \sum_i f_i \exp 2\pi i(hx_i+ky_i+lz_i) \right] \exp -2\pi i(hx+ky+lz)$$

oraz

$$\operatorname{tg}\varphi_{hkl} = B_{hkl} / A_{hkl}$$

Obok odgadnięcia struktury (przypadek praktycznie niemożliwy dla białek czy kwasów nukleinowych!) istnieje kilka eleganckich formalnie metod rozwiązania problemu fazowego.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Metoda Pattersona polega na zastosowaniu odwrotnej transformacji Fouriera do natężeń refleksów zmierzonych eksperymentalnie. Wyniki pomiaru należy przedtem skorygować o czynniki geometryczne i fizyczne wpływające na mierzone natężenie, a opisane poprawkami Lorentza  $L$ , polaryzacyjną  $p$  i absorpcji  $A$ :

$$I = LpA |F|^2 = LpA FF^*$$

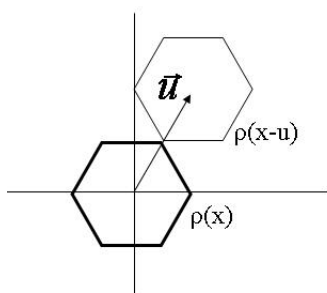
Odwołując się do wstępu matematycznego, można wartość transformaty w punkcie  $u,v,w$  zapisać:

$$P_{uvw} = V^{-1} \sum_{hkl} FF^* \exp(-2\pi i(hu + kv + lw))$$

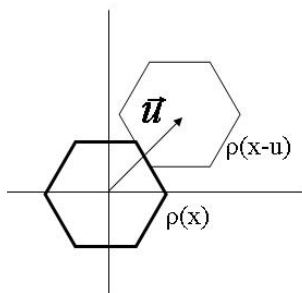
Zapis odpowiada splotowi funkcji rozkładu gęstości elektronowej  $\rho_{xyz}$  oraz funkcji przesuniętej o wektor  $[uvw]$

$$P_{uvw} = \int \rho_{xyz} \otimes \rho_{x-u, y-v, z-w} dV$$

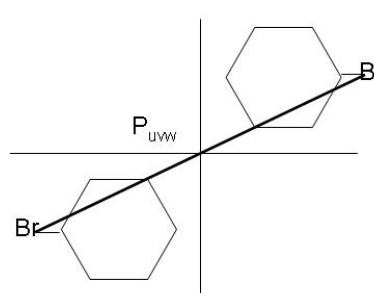
W punkcie  $uvw$  wartość funkcji Pattersona jest proporcjonalna do  $\rho_{xyz} \otimes \rho_{x-u, y-v, z-w}$ . Jeśli wektor  $uvw$  odpowiada wektorowi między atomami o dużej liczbie atomowej, gęstość elektronowa  $\rho$  związana z tymi atomami jest duża i wartość funkcji  $P_{uvw}$  jest duża. Jeżeli  $uvw$  nie jest wektorem międzyatomowym, to w  $(xyz)$  ( $x-u, y-v, z-w$ ) lub w obu punktach gęstość ma wartość bliską zero, zatem wartość funkcji Pattersona  $P$  jest mała lub zerowa (**Rys. 3.1** i **Rys. 3.2**).



**Rys. 3.1**



**Rys. 3.2**



**Rys. 3.3**

Jeśli  $uvw$  to składowe wektora między atomami, to  $P_{uvw} \sim l.a.(at1) \cdot l.a.(at2)$ , a składowe wektora są równe  $u=x_2-x_1$ ,  $v=y_2-y_1$ ,  $w=z_2-z_1$  (**Rys. 3.3**). Dla struktury bromobenzenu piki odpowiadające wektorom międzyatomowym Br-Br, Br-C i C-C będą miały wysokość proporcjonalną do 1225, 210 i 36. Dlatego po wyliczeniu wartości funkcji Pattersona łatwo byłoby zidentyfikować piki odpowiadające wektorom Br-Br. Trzeba zauważyć, że dla struktury  $N$ -atomowej mamy  $N(N-1)$  pików, a dodatkowo funkcja Pattersona wykazuje symetrię środka.



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Dalszym ułatwieniem interpretacji funkcji Pattersona są cięcia Harkera, definiowane jako wektory między atomami związanymi symetrią w danej grupie przestrzennej. Po zidentyfikowaniu grupy na podstawie symetrii obrazu dyfrakcyjnego i wygaszeń systematycznych można określić cięcia Harkera. Przykładowo dla grupy przestrzennej  $C2/c$  cięcia będą związane z działaniem sieci translacyjnej  $C$ , środka symetrii, osi dwukrotnej i płaszczyzny  $c$ . W tej grupie dla komórki  $C$  cięcie  $uvw$  odpowiada wektorowi  $\frac{1}{2}\frac{1}{2}0$ , dla środka  $2x,2y,2z$  (**Rys. 3.3**), dla osi dwukrotnej  $[010]$   $2x,0,2z$ , a dla płaszczyzny  $z$  poślizgiem  $c$  to wektor  $0,2y,\frac{1}{2}$ .

Pozostaje pytanie kiedy da się zastosować metodę Pattersona do rozwiązania problemu fazowego. Narzucająca się odpowiedź – dla struktur zawierających atomy cięższe od pozostałych. Jak ciężkie? To określa wzór Sima,  $\Sigma Z^2_{\text{heavy}} / \Sigma Z^2_{\text{light}} \sim 0.5$ . Jednak przy wartości ilorazu około 0.1 też można uzyskać rozwiązanie. Efektywnie można użyć tej metody dla rozwiązania struktury związków organicznych zawierających atom siarki lub cięższy. Obecność atomu Br praktycznie gwarantuje, że obliczone położenie atomu ciężkiego da fazy, które pozwolą obliczyć mapę gęstości elektronowej pokazującą cząsteczkę lub jej duży fragment. Natomiast metoda Pattersona nie da się zastosować do zwykłych związków organicznych, gdyż nie ma tam atomów dużo cięższych niż pozostałe - trudno odróżnić cięcia Harkera O-O (~64) od O-C (~48) i C-C (~36) gdy funkcja Pattersona ma  $N(N-1)$  pików.

Należy też podkreślić, że metoda Pattersona jest stosunkowo mało wrażliwa na błędy systematyczne wynikające z eksperymentu. Ważne jest uzyskanie rzeczywistych wartości  $|F|^2$  przez korektę zmierzonych natężeń refleksów o czynniki Lorentza, polaryzacyjny i absorpcję. Z drugiej strony konieczność zastosowania cięć Harkera powoduje, że trzeba poprawnie określić grupę przestrzenną.

Co zrobić, gdy nie można odgadnąć położenia atomów w strukturze (bo nie jest spełniony wymóg  $Z < \text{lprs}$ ) albo nie ma atomów istotnie cięższych od pozostałych? Dla struktur zawierających ‘jednakowo lekkie’ atomy stosuje się metody bezpośrednie, za opracowanie których Jerome Karle i Herbert Hauptmann uzyskali w 1985 roku nagrodę Nobla. Metody bezpośrednie oparte są na zależnościach między wartościami czynników struktury. Pierwszym etapem jest skalowanie refleksów i obliczenie ogólnego czynnika temperaturowego metodą Wilsona. Następnie prowadzi się normalizację czynników struktury zgodnie z równaniem, w którym licznik ułamka to uzyskane po skalowaniu kwadraty czynników struktury, mianownik zawiera sumę kwadratów atomowych czynników rozpraszania i mnożnik uwzględniający zmianę natężeń związaną z wygaszeniami systematycznymi:

$$E^2_{hkl} = |F_{hkl}|^2 / \Sigma f_i^2 \varepsilon$$

Dla transformat Fouriera  $G(h)$  gęstości elektronowej zachodzi związek w przestrzeni odwrotnej  $G(h) = \int G(h')G(h-h')dV^*$ . Podstawą teoretyczną jest wynikające z tego równanie Sayre’a, w którym całkę można zastąpić sumą po  $h'$  (refleksach)

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

$$E(h) = T \sum E(h')E(h-h')$$

Dla struktur centrosymetrycznych rozwiązanie jest względnie łatwe. Można wykazać, że dla nich faza może mieć tylko dwie wartości: 0 lub  $\pi$ . To odpowiada zależności

$$E(h) = |E(h)| \cos \alpha = |E(h)| s(h) \quad \text{gdzie } s(h) \text{ to } +1 \text{ lub } -1$$

Równanie Sayre'a dla dużych wartości znormalizowanych czynników struktury E dla trypletu refleksów o wskaźnikach h, h' oraz h-h' prowadzi do zależności  $\Sigma_2$ :

$$s(h) = s(h')s(h-h'),$$

której szczególnym przypadkiem jest zależność  $\Sigma_1$  gdy  $h = 2h'$  :  $s(h) = s(h')s(h')$ . Bez względu na wartość fazy refleksu h', faza refleksu h wynosi 0.

Prawdopodobieństwo określenia fazy z zależności  $\Sigma_2$  dla struktury N-atomowej:

$$p = 1/2 + 1/2 \operatorname{tgh} [1/N^{1/2} |E(h_1)E(h_2)E(h_3)|]$$

Prawdopodobieństwo rośnie dla refleksów o dużych  $E(h_1)$ ,  $E(h_2)$ ,  $E(h_3)$ , maleje dla dużych N. Wynika z tego, że metody bezpośrednie można stosować dla stosunkowo małych struktur, ale nie dla białek.

Dla struktur niecentrosymetrycznych nie ma ograniczenia wartości fazy refleksów. Stosuje się wzór tangensowy Karle-Hauptmanna

$$\operatorname{tg} \alpha(h) = \frac{\sum |E(h')E(h-h') \sin[\alpha(h) + \alpha(h-h')]}{\sum |E(h')E(h-h') \cos[\alpha(h) + \alpha(h-h')]}$$

Ta metoda obliczenia fazy refleksu hkl może być stosowana dla struktur związków organicznych i metaloorganicznych. Należy pamiętać, że u podstaw metod bezpośrednich leży założenie o identyczności wszystkich atomów w strukturze i ich jednakowych oscylacjach wokół położenia równowagi. Jeśli to uproszczenie jest zbyt drastyczne, metoda może zawieść. Kolejnym ograniczeniem jest normalizacja, wymagająca znajomości składu związku w badanym kryształ. Wprowadzenie błędów na tym etapie może utrudnić uniemożliwić rozwiązanie problemu fazowego. Z kolei niepoprawna symetria grupy przestrzennej może spowodować uśrednienie wartości czynników struktury nie związanych symetrią w rzeczywistym kryształ, co także może uniemożliwić rozwiązanie struktury.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
 realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Znormalizowane czynniki struktury  $E_{hkl}$  mają takie same właściwości jak czynniki struktury  $F_{hkl}$ . Odwrotna transformacja Fouriera prowadzi do wyliczenia map rozkładu gęstości elektronowej. Ze względu na rodzaj zastosowanych czynników struktury, mamy kilka rodzajów funkcji opisujących rozkład gęstości elektronowej. We wszystkich przypadkach fazy  $\varphi$  są obliczane w oparciu o współrzędne zlokalizowanych atomów (metody bezpośrednie, metoda Pattersona, odgadnięcie).

Funkcja  $\rho_{calc}$  opisująca obliczony rozkład gęstości elektronowej odpowiada aktualnemu modelowi rzeczywistej struktury kryształu. Jest ona liczona jako odwrotna transformata Fouriera obliczonych czynników struktury  $F_{c,hkl}$  będących sumą wkładów atomów o znanych współrzędnych:

$$F_{c,hkl} = \sum f [\cos 2\pi (hx+ky+lz) + i \sin 2\pi (hx+ky+lz)] = A + iB$$

$$\rho_{calc} = V^{-1} \sum_{hkl} F_{c,hkl} \exp -2\pi i (hx+ky+lz)$$

Czynnik  $F_{o,hkl}$  ( $F_{hkl}$  obserwowany) obliczany jest z użyciem  $|F_{hkl}|$  z eksperymentu oraz fazy  $\varphi_c$  określonej z częściowo znanej struktury:

$$F_o = |F| \exp i\varphi_c \quad \text{gdzie} \quad \varphi_c = \text{tg}^{-1} [\sum f \sin 2\pi (hx+ky+lz) / \sum f \cos 2\pi (hx+ky+lz)]$$

Obserwowana gęstość elektronowa  $\rho_{obs}$  jest liczona jako odwrotna transformata  $F_o$ . Jest to mapa rzeczywistego rozkładu gęstości eksperymentalnej, tej, która powoduje powstanie zmierzonego obrazu dyfrakcyjnego kryształu.

$$\rho_{obs} = V^{-1} \sum_{hkl} F_o \exp -2\pi i (hx+ky+lz)$$

Bardzo ważna i użyteczna jest transformata z różnicy funkcji  $F_o$  i  $F_c$ , odpowiadająca różnicy funkcji rozkładu gęstości  $\rho_{obs} - \rho_{calc} = \Delta\rho$ . Mapa ta nazywana jest mapą różnicową. Różnica czynników struktury refleksów  $hkl$  liczona jest jako:

$$F_o - F_c = |F| \exp i\varphi_c - \sum f_i \exp 2\pi i (hx_i + ky_i + lz_i)$$

gdzie faza  $\varphi_c = \text{tg}^{-1} [\sum f \sin 2\pi (hx+ky+lz) / \sum f \cos 2\pi (hx+ky+lz)]$

$$\Delta\rho(xyz) = V^{-1} \sum_{hkl} (F_o - F_c) \exp -2\pi i (hx+ky+lz) = \rho_{obs}(xyz) - \rho_{calc}(xyz)$$

Funkcja różnicowa służy do kompletowania i korygowania błędów struktury. Wartość  $\Delta\rho(xyz) < 0$  oznacza, że  $\rho_{obs} < \rho_{calc}$ , czyli w modelu w położeniu  $xyz$  mamy za wysoką wartość  $\rho_{calc}$  w porównaniu z rzeczywistą gęstością – za ciężki atom lub atom w błędnym położeniu. Jeśli  $\Delta\rho(xyz)$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

$> 0$  oznacza, że  $\rho_{\text{obs}} > \rho_{\text{calc}}$ . To z kolei znaczy, że w punkcie xyz brak atomu lub w modelu znajduje się tam atom o zbyt małej liczbie atomowej. Interpretacja map różnicowych gęstości pozwala na skorygowanie niepoprawnej konformacji cząsteczki białka, a także wskazuje na obecność rozpuszczalnika w strukturze kryształu.

Do oceny jakości rozwiązania struktury używa się tzw. wskaźnika rozbieżności R:

$$R = \frac{\sum \|F_o\| - \|F_c\|}{\sum \|F_o\|}$$

Dla początkowych etapów budowy modelu  $R \sim 30\%$ . Dla dobrze udokładnionych struktur małowcząsteczkowych  $R \sim 2-6\%$ , dla dobrze określonych struktur białek wskaźnik rozbieżności  $R \sim 10-20\%$ .

## Wykład 5: Metody proteomiki strukturalnej

Proteomika strukturalna posługuje się kilkoma metodami pozwalającymi na określenie struktury przestrzennej białek. Należą do nich krystalografia rentgenowska lub stosująca inne rodzaje promieniowania (wiązki elektronów i neutronów), spektroskopia magnetycznego rezonansu jądrowego NMR, oraz mikroskopia elektronowa EM. Ze względu na dostępność metody i wagę wyników, najczęściej stosowana jest krystalografia rentgenowska. W banku struktur białkowych i kwasów nukleinowych PDB, na 96692 zdeponowanych struktur, metodami krystalograficznymi rozwiązano 85486 struktur (88.4 %), 10.6 % struktur określono metodą NMR, a tylko 716 struktur metodą mikroskopii elektronowej. Innymi metodami określono 176 struktur, w tym 44 neutronograficznie i 38 elektronograficznie. Te dane (01.01.2014) jednoznacznie wskazują na wiodącą rolę krystalografii rentgenowskiej w badaniach biomakrocząsteczek.

W krystalografii strukturalnej najważniejszy jest problem fazowy. Aby obliczyć mapę rozkładu materii, potrzebujemy zarówno amplitud czynników struktury, jak i ich faz. Obliczenie czynników struktury  $F_{hkl} = |F_{hkl}| \exp(i\phi)$  umożliwia obliczenie odwrotnej transformaty Fouriera

$$\rho_{xyz} = V^{-1} \sum_{hkl} F_{hkl} \exp(-2\pi i(hx+ky+lz))$$

Fazę można wyliczyć ze znanych położenia atomów w komórce  $x_i, y_i, z_i$ . Dla struktur małowcząsteczkowych zawierających ciężki atom, można zastosować metodę Pattersona posługując się cięciami Harkera. Jednak stosowalność tej metody ograniczona jest warunkiem Sima:

$$\sum Z_{\text{heavy}}^2 / \sum Z_{\text{light}}^2 \sim 0.5$$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Rozważenie składu białka zbudowanego z naturalnie występujących aminokwasów wskazuje, że liczba atomów siarki (Cys, Met) w typowym białku nie jest wystarczająca dla rozwiązania problemu fazowego w klasyczny sposób. Także w przypadku metaloenzymów wkład jonów metalu nie będzie nawet w przybliżeniu spełniał tego warunku. Z tego punktu widzenia cząsteczkę białka należałoby traktować tak jak związki organiczne. Przykładem może być NTPDaza1 z *S. tuberosum*, dla której w 455-aminokwasowej sekwencji jest 8 cystein i 7 metionin. Jeśli atomy S traktować jako atomy ciężkie, to iloraz  $S_{\text{ima}}$  przyjmuje wartość 0.0595

Można próbować zastosować metody bezpośrednie do otrzymania zbioru faz i dalszego obliczenia funkcji rozkładu gęstości elektronowej. Prawdopodobieństwo przypisania fazy w oparciu o zależność  $\Sigma_2$  dla struktury N-atomowej rośnie dla refleksów o dużych  $E(h_1)$ ,  $E(h_2)$ ,  $E(h_3)$ , maleje dla dużych struktur:

$$p = 1/2 + 1/2 \operatorname{tgh} [1/N^{1/2} | E(h_1)E(h_2)E(h_3) | ]$$

Dla białka zbudowanego z 500 aminokwasów przeciętnie zawierających 10 atomów niewodorowych prawdopodobieństwo będzie zależeć od  $1/70$  iloczynu trzech wartości  $E$ , co da wartość prawdopodobieństwa zbliżoną do  $1/2$ . Użycie takiej fazy do dalszych zależności  $\Sigma_2$  niemożliwi obliczenie poprawnych map rozkładu gęstości elektronowej. Jak więc radzą sobie wszyscy autorzy zdeponowanych struktur?

Jedną z metod jest **metoda podstawienia cząsteczkowego MR** (Molecular Replacement). Białko zbudowane jest z aminokwasów połączonych w łańcuch polipeptydowy. Łańcuchy boczne aminokwasów tworzą oddziaływania z otaczającymi fragmentami białka, przy czym oddziaływania te zależą od rodzaju aminokwasu. Także grupy łańcucha głównego tworzą sieć wiązań wodorowych z otoczeniem. Te oddziaływania powodują zmiany konformacyjne łańcucha głównego i przyjęcie przez polipeptyd optymalnej konformacji odpowiadającej minimum energetycznemu. Można przyjąć, że w jednakowych warunkach fizyko-chemicznych łańcuch polipeptydowy o określonej sekwencji aminokwasowej osiągnie równowagę termodynamiczną przyjmując zawsze taką samą konformację, odpowiadającą globalnemu minimum energii.

Jeśli badamy nieznaną białko, możemy poszukać białka o jak najbardziej podobnej sekwencji, ale o znanej strukturze przestrzennej. To białko będzie modelem białka nieznanego. Jeśli identyczność sekwencji będzie wysoka, można założyć, że sfałdowanie białka modelowego będzie bardzo zbliżone do białka badanego. Wróćmy do funkcji Pattersona. Każdą cząsteczkę można przedstawić jako kompletny zbiór wektorów między jej atomami. Na tej podstawie można wyliczyć wartości funkcji Pattersona odpowiadające cząsteczce białka. Funkcja Pattersona wyliczona na podstawie struktury białka modelowego  $P_{uvw,model}$  powinna więc odpowiadać funkcji Pattersona białka badanego. W eksperymencie dyfrakcyjnym mierzymy natężenia refleksów, które

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

przeliczamy na kwadraty amplitud  $|F_{hkl}|^2$  i znaną metodą obliczamy transformatę Fouriera uzyskując eksperymentalną funkcję Pattersona dla nieznanego białka  $P_{uvw,exp}$ .

Procedurę dopasowania obu funkcji Pattersona można podzielić na dwa etapy. W pierwszym dokonujemy obrotu modelowej cząsteczki białka z niewielkim krokiem kątowym, aby znaleźć najlepsze dopasowanie – posługujemy się tzw. funkcją rotacyjną RF. Prawidłowa orientacja cząsteczki modelowej powinna odpowiadać orientacji cząsteczki badanej w komórce elementarnej. Należy zaznaczyć, że to najdłuższe wektory międzyatomowe pozwalają na dokładne orientowanie cząsteczki, ale muszą to być wektory wewnątrzcząsteczkowe.

$$R(\phi, \psi, \kappa) = \int P_{exp}(u) P_{model}(\phi, \psi, \kappa, u) du$$

W drugim etapie, po znalezieniu właściwej orientacji, poszukuje się położenia tak orientowanej cząsteczki w komórce elementarnej badając funkcję translacyjną TF odpowiadającą wektorom międzycząsteczkowym. Można to robić badając korelację pomiędzy funkcjami Pattersona lub czynnikami struktury, jak w równaniu:

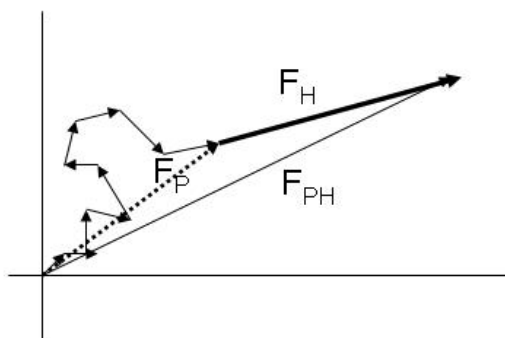
$$T(r) = \int F_{exp}(r) F_{model}(\phi, \psi, \kappa)(r) dr$$

Maksimum funkcji translacyjnej odpowiada prawdziwemu położeniu cząsteczki o orientacji określonej w pierwszym kroku. Zwraca uwagę fakt, że przeszukiwanie funkcji translacyjnej należy dokonywać w gęstej siatce punktów xyz, aby nie pominąć prawdziwego rozwiązania. To wymaga każdorazowego wyliczenia czynników struktury  $F_{hkl}$  dla przesuniętej cząsteczki próbnej i jest to duży problem obliczeniowy.

Rodzi się pytanie, jak podobna ma być cząsteczka modelowa do cząsteczki badanej. Prowadzono badania porównawcze struktur białek i niezależnie uzyskanych modeli cząsteczek, wiążąc zmiany konformacji białka ze zmianami w jego sekwencji. Wyniki prowadziły do wniosku, że 30 % identyczności sekwencji odpowiada średnia kwadratowa odległość położenia atomów C $\alpha$  od rzeczywistej około 1.5 Å. Ta wartość wydaje się dobrą wskazówką stosowalności modelu. Przy większej różnicy położenia atomów C $\alpha$  rmsd > 1.5 Å mogą wystąpić trudności ze znalezieniem maksimum RF i TF. Jednak identyczność sekwencji na poziomie 20% też daje szansę na sukces przy zastosowaniu metody MR.

Teoria czynników struktury podpowiada kolejną metodę. Jest nią **metoda podstawienia izomorficznego MIR** (Multiple Isomorphous Replacement). Addytywności wkładów atomowych powoduje, że czynnik struktury można przedstawić jako sumę wektorowych wkładów atomowych (**Rys. 5.1**)

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki



Rys. 5.1

Suma wkładów atomów białka daje wypadkowy wektor  $F_P$ , odpowiadający czynnikowi struktury natywnego białka. Załóżmy, że białko wiąże ciężki atom, co daje w rezultacie czynnik struktury kompleksu białko-ciężki atom  $F_{PH}$ . Obie wartości różnią się wkładem ciężkiego atomu  $F_H$  do czynnika struktury, który zgodnie z teorią jest znacznie większy od każdego z pozostałych atomów. Jeśli ocenimy wkład ciężkiego atomu przez iloraz Sima, na pewno nie przekroczymy krytycznej wartości pozwalającej na bezpośrednie zastosowanie metody Pattersona.

Wiadomo, że białko stanowi 20-80 % zawartości kryształów, resztę stanowi rozpuszczalnik, tworzący tzw. kanały wodne. To umożliwi wprowadzenie ciężkich atomów przez zanurzenie kryształu białka w roztworze soli ciężkich jonów, czyli metodą nasączania. Jony ciężkie będą związane zgodnie z ich powinowactwem do właściwych chemicznie grup, często na powierzchni białka, nie powodując zmian konformacji. Białko i kompleks mają tę samą strukturę – mówimy, że są izomorficzne. To spełnia wymóg, że czynnik struktury natywnego białka  $F_P$  jest identyczny z wkładem atomów białka w czynniku struktury kompleksu  $F_{PH}$ .

Można zatem przeprowadzić dwa eksperymenty dyfrakcyjne, mierząc w pierwszym  $I_{hkl,P}$  dla kryształu natywnego białka, a w drugim  $I_{hkl,PH}$  dla kompleksu. Przy zachowaniu izomorficzności obu struktur, jedyna różnica powodowana jest przez wkład  $F_H$ . Obliczamy różnicę

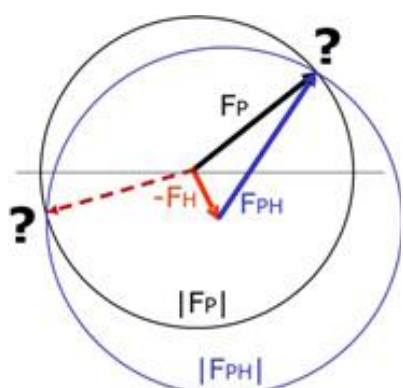
$$|F_{hkl,PH}|^2 - |F_{hkl,P}|^2 = |F_{hkl,H}|^2$$

Różnicowy kwadrat amplitudy ma 100% wkładu ciężkiego atomu! To pozwala metodą cięć Harkera zlokalizować ten atom i wyliczyć  $F_H$ . Aby uzyskać taką informację trzeba dokonać pomiaru dla dwóch kryształów izomorficznych – natywnego białka i jego kompleksu. Niestety często otrzymanie ciężkiej pochodnej powoduje brak izomorficzności i różnice kwadratów amplitudy są wynikiem obecności ciężkich atomów w nieznanym położeniu i jednocześnie nieznanymi zmianami położenia atomów białka. Teraz problem fazowy sprowadza się do rozwiązania wektorowego równania odpowiadającego Rys. 5.1:

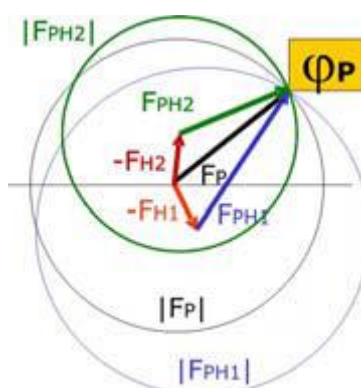
$$F_P = F_{PH} - F_H$$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Jest to podstawa metody SIR (Single Isomorphous Replacement), a równanie można przedstawić graficznie (**Rys. 5.2**). Z początku układu rysujemy okrąg o promieniu  $|F_P|$ . Następnie ze środka okręgu prowadzimy wektor  $-F_H$ , a z tego punktu rysujemy okrąg o promieniu  $|F_{PH}|$ . Przecięcie obu okręgów oznacza rozwiązanie równania wektorowego. Jednak równanie ma dwa rozwiązania, a więc i dwie fazy dla refleksu  $F_{hkl,P}$ . Takie rozwiązanie należy przeprowadzić dla wszystkich refleksów. Można oczywiście próbować wybrać jedną z faz i wyliczyć mapy rozkładu gęstości  $\rho_{xyz}$ . Dla eksperymentu z 50000 refleksów mamy do sprawdzenia tylko  $2^{50000}$  możliwych rozwiązań, a jedno jest prawdziwe! Eksperymenty dyfrakcyjne z zastosowaniem źródła synchrotronowego często dają 1 milion refleksów.



**Rys. 5.2**



**Rys. 5.3**

Opisaną powyżej niejednoznaczność można łatwo rozwiązać otrzymując kolejną ciężką pochodną białka. Musi ona być izomorficzna z białkiem natywnym, aby także dla niej można obliczyć różnice natężeń refleksów jak dla pochodnej pierwszej. Teraz dla każdego refleksu  $hkl$  formułujemy układ równań wektorowych analogicznych do pierwszego, a jego rozwiązanie ilustruje **Rys. 5.3**. Użycie kilku ciężkich pochodnych stanowi istotę metody MIR. W identycznej procedurze lokalizujemy ciężkie atomy w obu pochodnych, a rozwiązanie dwóch równań jednoznacznie określa fazę czynnika  $F_P$  dla białka!. Mając czynniki dla białka można wyliczyć wartość funkcji rozkładu gęstości elektronowej  $\rho_{xyz}$ . Metoda MIR jest skuteczną metodą rozwiązania problemu fazowego dla biomakrocząsteczek. Można w niej zastosować dowolną liczbę pochodnych, wprowadzając ciężkie atomy Hg, U, Au, Os, Ir, Pt, Pd czy inne, jednak każda z nich musi być izomorficzna z białkiem natywnym, a to zwiększa z każdym kryształem ryzyko wystąpienia braku izomorficzności.

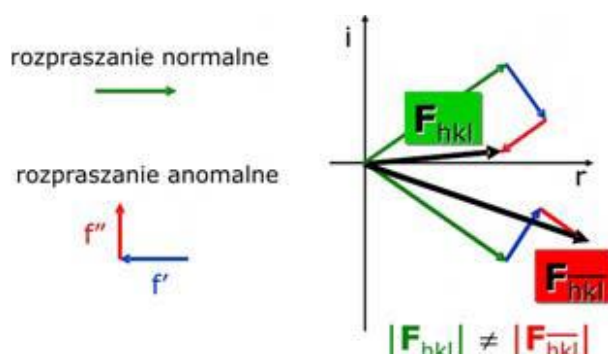
Czy można użyć jednego kryształu, aby wykorzystać technikę opisaną powyżej? Odpowiedź jest pozytywna. Musimy jednak wrócić do teorii atomowych czynników rozpraszania. W zwykłych eksperymentach atomowy czynnik rozpraszania jest proporcjonalny do  $f_0$  – czynnika dla atomu nie wykonującego w sieci oscylacji wokół położenia równowagi. Ten z kolei jest proporcjonalny do liczby atomowej danego pierwiastka. Jednak w pobliżu progu absorpcji atomowy czynnik rozpraszania można zapisać jako liczbę zespoloną:

$$f_a = f_0 + f' + if''$$



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Pierwszy składnik sumy  $f_0$  to część normalna rozpraszania atomowego, suma  $f'+if''$  to tzw. poprawka anomalna. Jej część rzeczywista  $f'$  jest zwykle ujemna, czyli obniża wartość atomowego czynnika rozpraszania. Część urojona  $if''$  (poprawka absorpcyjna) ma wartość wprost proporcjonalną do współczynnika absorpcji. W układzie osi (rzeczywistej i urojonej) obie poprawki przedstawić można jako parę wektorów, rzeczywistego  $f'$  oraz prostopadłego do niego urojonego  $f''$ . Anomalne rozpraszanie powoduje złamanie prawa Friedla  $F_{hkl} \neq F_{-h-k-l}$ . Dla refleksów o przeciwnych znakach części rzeczywiste mają przeciwne fazy, czyli są wektorami położonymi symetrycznie względem osi rzeczywistej. Jednak ze względu na stały kąt pomiędzy wektorami opisującymi składową anomalną, części  $f''$  obu wiązek powodują różnicę (**Rys. 5.4**).



Rys. 5.4

Poprawka anomalna zależy od długości fali promieniowania. Dysponując źródłem promieniowania o regulowanej długości fali, można dostroić promieniowanie aby anomalne rozpraszanie było maksymalne dla danego pierwiastka. Stąd bierze się nazwa – **metoda dostrojonego rozpraszania anomalnego MAD** (Multiwavelength Anomalous Diffraction). Przy dostrojonej długości fali, pozostałe pierwiastki nie wykazują anomalnego rozpraszania.

Historycznie, rozpraszanie anomalne wykorzystywano rejestrując wiązki  $hkl$  i  $-h-k-l$ , a z obliczonej różnicy natężeń lokalizowano atomy rozpraszające anomalnie. Obecnie zastosowanie źródeł synchrotronowych pozwala dla jednego kryształu wykonać pomiar przy trzech długościach fali, optymalnych dla  $f'$ ,  $f''$  oraz takiej, gdzie efekt jest mniejszy. To odpowiednik dwóch pochodnych i kryształu natywnego z metody MIR. Istotne jest, że użycie jednego kryształu wyklucza brak izomorficzności.

Dalsze postępowanie jest analogiczne dla opisanego w metodzie MIR. Obliczenie różnicowych funkcji Pattersona  $\Delta|F_{hkl}|^2$  pozwala standardową metodą cięć Harkera lub metodami bezpośrednimi zlokalizować położenie atomu dającego anomalne rozpraszanie i wyliczyć jego wkład  $F_{anom}$ . Rozwiązanie problemu fazy jest identyczne jak w MIR. Można ograniczyć się do analogu metody SIR, wtedy mówimy o pomiarze przy dwóch długościach fal i metodzie SAD (Single Anomalous Diffraction)

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Metoda MAD jest najbardziej zaawansowaną teoretycznie i technicznie metodą krystalografii i stała się potężnym narzędziem proteomiki strukturalnej. Dzięki niej można rozwiązać struktury białek nie mających homologów wśród dotychczas zbadanych białek zdeponowanych w PDB.

## Wykład 6: Elementy bioinformatyki – modelowanie molekularne

Bioinformatykę można określić jako zastosowanie metod komputerowych w rozwiązywaniu problemów biologicznych. Stosowalność i możliwości bioinformatyki nieustannie rosną wraz ze wzrostem mocy obliczeniowej komputerów, rozwojem internetu oraz rozwojem baz danych biologicznych – w tym baz sekwencji, motywów strukturalnych i struktur biomakrocząsteczek.

Istotna część badań bioinformatycznych to przewidywanie struktury drugorzędowej białka na podstawie analizy pojedynczej sekwencji. Metoda oparta jest na analizie częstości występowania aminokwasów w elementach drugorzędowej struktury takich jak helisy, wstęgi  $\beta$  czy skręty  $\beta$ , a także na właściwościach aminokwasów w tym Gly, Pro i Asn. Połączenie analizy sekwencji z analizą hydrofobowości pozwala z pewnym prawdopodobieństwem przewidywać konformację fragmentów łańcucha polipeptydowego.

Bazy sekwencji zawierają zarówno sekwencje genów (GenBank, DDBJ, EMBL), sekwencji białkowych (SWISS-PROT, PIR), struktur białek (PDB) jak i modeli cząsteczek białek (MMDB). Najbardziej podstawowy sposób wykorzystania tych baz sekwencji to identyfikacja swojego obiektu badań lub makrocząsteczek pokrewnych przez znalezienia sekwencji identycznych lub bardzo zbliżonych. Stosuje się algorytmy porównujące dwie lub wiele sekwencji. Przeszukanie baz sekwencji nukleotydowych jest stosunkowo proste ze względu na ograniczoną liczbę możliwości (5 nukleotydów). Znacznie bardziej złożona jest analiza sekwencji białek, gdyż zbudowane są one z 20 naturalnie występujących aminokwasów. Aminokwasy można grupować według hydrofobowości, obecności grup polarnych i naładowanych w łańcuchach bocznych, co sugeruje podobieństwo aminokwasów w danej grupie.

W toku ewolucji występują zmiany w sekwencji aminokwasowej. Podobieństwo aminokwasów sugeruje, że można wprowadzić do sekwencji punktową mutację aminokwasu bez utraty funkcjonalności białka. Jednak mutacje punktowe mogą dotyczyć zamiany na aminokwas nie wykazujący podobieństwa, zmieniając aktywność białka. Do porównania sekwencji aminokwasowych potrzebne są algorytmy, które wskażą sekwencje podobne lub identyczne, z jednoczesnym uwzględnieniem różnic długości porównywanych sekwencji i różnic aminokwasów na danej pozycji w sekwencji. Aby prowadzić porównanie sekwencji białek, stosuje się macierze prawdopodobieństwa. Dwie najbardziej znane to macierze PAM i BLOSUM.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Macierze PAM (Point Accepted Mutation) zostały opracowane przez Margaret Dayhoff. Zależą one od ewolucyjnego dystansu między parami aminokwasów. Tworzenie macierzy można podzielić na kilka etapów. Początkowo tworzy się mapę podobieństw ewolucyjnych wśród badanych sekwencji białkowych i oblicza częstość zmian aminokwasów na inny. Następnie dla standardowej charakterystycznej sekwencji analizuje się częstość podstawień nie zmieniających funkcji białka (częstość docelowa) i częstość występowania różnych aminokwasów (częstość tła). Stąd PAM to macierz 20×20 uwzględniająca tempo i typ możliwych podstawień.

$$M_{ij} = P_{ij}(\delta t) = \lambda \frac{A_{ij}}{N_i} \quad \text{gdzie}$$

- $P_{ij}(\delta t)$  - prawdopodobieństwo zamiany aminokwasu  $i$  na  $j$  w czasie  $\delta t$
- $\delta t$  - tempo zmian aminokwasów
- $N_i$  - częstość zmian „przypadkowych” – częstość tła (liczba wystąpień aminokwasu  $i$  w całym zbiorze sekwencji)
- $A_{ij}$  - częstość substytucji – częstość docelowa – zmiany pojawiające się w białkach spokrewnionych (liczba obserwowanych podstawień)

Elementy macierzy  $Y_{ij}$  (**Rys. 6.1**) oblicza się jako logarytm naturalny  $\ln M_{ij}$ , mnoży przez 10 i zaokrągla do wartości całkowitych. Jedna jednostka PAM oznacza stopień zróżnicowania ewolucyjnego, w którym zmienił się 1% aminokwasów. Ekstrapolacja dla różnych odległości PAM pozwala tworzyć różne macierze. PAM250 jest macierzą dla odległych ewolucyjnie sekwencji i może służyć do porównania sekwencji białek należących do różnych grup. Zatem macierz PAM40 jest pozwala na porównanie sekwencji białek bliskich (40 jednostek PAM), zaś PAM250 jest dobra do przeszukiwania ogólnych białkowych baz danych.

A	2																			
R	-2	6																		
N	0	0	2																	
D	0	-1	2	4																
C	-2	-4	-4	-5	12															
Q	0	1	1	2	-5	4														
E	0	-1	1	3	-5	2	4													
G	1	-3	0	1	-3	-1	0	5												
H	-1	2	2	1	-3	3	1	-2	6											
I	-1	-2	-2	-2	-2	-2	-3	-2	5											
L	-2	-3	-3	-4	-6	-2	-3	-4	-2	2	6									
K	-1	3	1	0	-5	1	0	-2	0	-2	3	5								
M	-1	0	-2	-3	-5	-1	-2	-3	-2	2	4	0	6							
F	-3	-4	-3	-6	-4	-5	-5	-5	-2	1	2	-5	0	9						
P	1	0	0	-1	-3	0	-1	0	0	-2	-3	-1	-2	-5	6					
S	1	0	1	0	0	-1	0	1	-1	-1	-3	0	-2	-3	1	2				
T	1	-1	0	0	-2	-1	0	0	-1	0	-2	0	-1	-3	0	1	3			
W	-6	2	-4	-7	-8	-5	-7	-7	-3	-5	-2	-3	-4	0	-6	-2	-5	17		
Y	-3	-4	-2	-4	0	-4	-4	-5	0	-1	-1	-4	-2	7	-5	-3	-3	0	10	
V	0	-2	-2	-2	-2	-2	-1	-2	4	2	-2	2	-1	-1	-1	0	-6	-2	4	
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V

Rys. 6.1 Macierz PAM250

A	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
C	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													
G	0	-2	0	-1	-3	-2	-2	6												
H	-2	0	1	-1	-3	0	0	-2	8											
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
T	0	-1	0	-1	-1	-1	-2	-2	-1	-1	-1	-2	-1	-1	1	5				
W	-3	-3	-4	-4	-2	-3	-2	-3	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Y	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7		
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V

Rys. 6.2 Macierz BLOSUM62

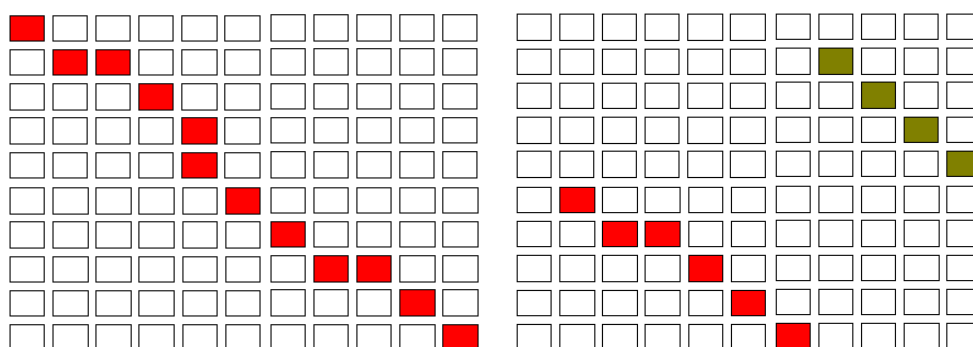
Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Macierze BLOSUM (Blocks Substitution Matrix) (**Rys. 6.2**) konstruuje się podobnie. Zlicza się częstości  $A_{ij}$  występowania aminokwasu  $i$  w danym miejscu łańcucha oraz aminokwasu  $j$  w tym samym miejscu w innych łańcuchach, szacuje częstość występowania aminokwasu, a w końcu oblicza funkcję oceny, przybliżając do najbliższej liczby całkowitej. Macierze BLOSUM opisuje się podając procent identyczności sekwencji. Stąd macierz BLOSUM62 została obliczona dla sekwencji o identyczności 62%, BLOSUM30 dla białek odległych ewolucyjnie (niski stopień identyczności), a BLOSUM90 do porównania sekwencji blisko spokrewnionych

Macierze BLOSUM bazują tylko na danych doświadczalnych, uwzględniają tylko rejony bez przerw (usunięć i wstawień) oraz dotyczą liczby dopasowanych par, a nie liczby podstawień. Z kolei PAM bazuje na blisko spokrewnionych białkach, wyniki ekstrapoluje się na inne białka. Wyższe numery w macierzach PAM oznaczają większy dystans ewolucyjny, natomiast w macierzach BLOSUM oznaczają mniejszy dystans ewolucyjny. Na podstawie sposobu tworzenia macierzy PAM i BLOSUM można sugerować odpowiedniość obu typów macierzy:

- PAM100 ==> Blosum90
- PAM120 ==> Blosum80
- PAM160 ==> Blosum60
- PAM200 ==> Blosum52
- PAM250 ==> Blosum45

Mając macierze podobieństw można prowadzić dopasowanie sekwencji. Dopasowanie globalne obejmuje pełne sekwencje, dopasowanie lokalne dotyczy fragmentów wykazujących najwyższe podobieństwo (**Rys. 6.3**)



**Rys. 6.3. Dopasowania globalne i lokalne.**

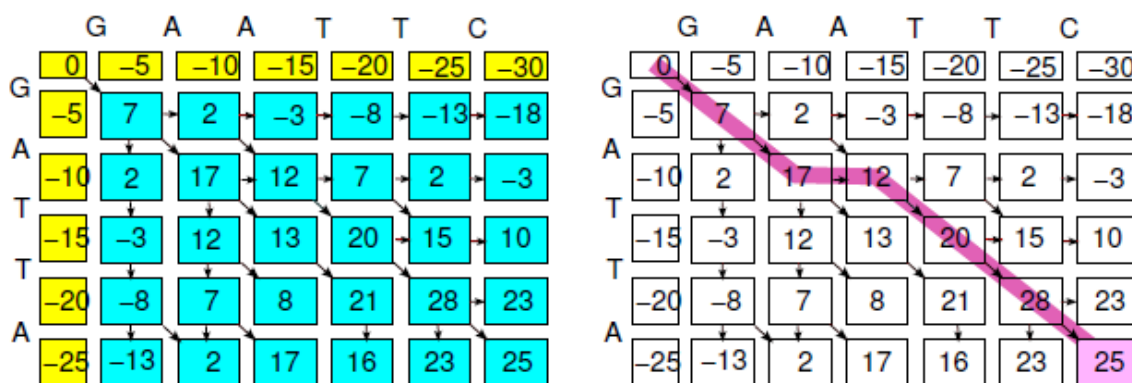
Ogólnie dopasowanie sekwencji prowadzi się stosując ukryte modele Markowa, w których prawdopodobieństwo wystąpienia sekwencji  $ab$  (prawdopodobieństwo warunkowe) zależy tylko od prawdopodobieństwa wystąpienia aminokwasu  $a$  w pozycji  $i-1$  i aminokwasu  $b$  w pozycji  $i$ , lecz jest niezależne od prawdopodobieństwa wystąpienia innych aminokwasów w pozycjach poprzedzających.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

$$P_{ab} = P(x_i=b \mid x_{i-1}=a) = P(x_i=b \mid x_{i-1}=a, x_{i-2}=m, x_{i-3}=n, x_{i-4}=0\dots)$$

Do dopasowania globalnego stosuje się algorytm Needlemana – Wunscha, który pozwala znaleźć maksymalną wartość oceny dopasowania  $\Sigma F(i,j)$  oraz sekwencję dla maksymalnej wartości dopasowania. Aby dopasować sekwencje o różnej długości wprowadza się karę  $d$  (ujemny wynik) za przerwy. Wartość funkcji dopasowania oblicza się jako sumę wartości z macierzy podobieństw liczoną w kierunkach wskazanych strzałkami na **Rys. 6.4**. Punkt startu to  $F(0,0)=0$ , koniec to  $(m,n)$ . Pokazane dopasowanie odpowiada sekwencji GA-TTC z przerwą w pozycji 3. Funkcje oceny  $F(i-1,j)$  oraz  $F(i,j-1)$  odpowiadają przerwom w jednej z dopasowywanych sekwencji.

$$F(i,j) = \max \begin{cases} F(i-1, j-1) + e(s_i, t_j) \\ F(i-1, j) + d \\ F(i, j-1) + d \end{cases}$$



**Fig 6.4** Dopasowanie globalne sekwencji GATTA do sekwencji GAATTC.

Dla dopasowania lokalnego stosuje się algorytm Smitha-Watermana. W tym algorytmie start  $F(i,0) = 0, F(0,j) = 0$ . Wartość funkcji oceny  $F(i,j)$  oblicza się jako

$$F(i,j) = \max \begin{cases} F(i-1, j-1) + e(s_i, t_j) \\ F(i-1, j) + d \\ F(i, j-1) + d \\ 0 \end{cases}$$

Najczęściej stosowane programy do dopasowania sekwencji należą do rodziny BLAST. Pozwalają one porównać i dopasować nie tylko sekwencje białek czy kwasów nukleinowych, ale także prowadzić porównania mieszane, np. sekwencji białka i kwasów nukleinowych (trójki kodujące). W zależności od wyniku dopasowania można zidentyfikować białko (identyczność sekwencji), określić przynależność badanego białka do rodziny znanych białek oraz funkcję i

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

architekturę (wysoki procent identyczności), albo zidentyfikować podjednostki strukturalne czy motywy strukturalne o znanej funkcji (podobieństwo fragmentów sekwencji).

Jeżeli stwierdzi się dopasowanie sekwencji białka o nieznannej strukturze do białka o znanej strukturze przestrzennej zdeponowanej w banku PDB lub MMDB, można zbudować przestrzenny model białka posługując się strukturą z banku jako szablonem. Taki proces nazywamy modelowaniem przez homologię (*homology modeling*). Kryterium ilościowe dla użyteczności struktury przestrzennej z banków PDB/MMDB jest co najmniej 30% identyczności sekwencji, bo w takim przypadku można oczekiwać podobnego sfałdowania białka (patrz Wykład 5, metoda MR).

Modelowanie wymaga zamiany aminokwasów na właściwe dla badanej sekwencji. To powoduje tworzenie luk w strukturze przestrzennej albo powstanie oddziaływań odpychających związanych z zatłoczeniem (zamiana dużego łańcucha bocznego na mniejszy lub odwrotnie). Konieczne jest też dopasowanie konformacji łańcucha polipeptydowego i łańcuchów bocznych. Otrzymanie optymalnej struktury wymaga więc minimalizacji energii całkowitej układu (cząsteczki) przez zmiany konformacyjne z zachowaniem geometrii walencyjnej. Dla dużych układów takich jak białka, stosuje się metody Mechaniki Molekularnej MM i Dynamiki Molekularnej MD.

W metodzie MM cząsteczkę traktuje się jak zbiór atomów połączonych ze sobą oddziaływaniami o określonych stałych siłowych. Całą cząsteczkę można opisać przez zbiór odległości międzyatomowych odpowiadających wiązaniom chemicznym (odległości 1-2), kątom walencyjnym (odległości 1-3), kątom torsyjnym (odległości 1-4) a także oddziaływaniami niewiążącym pomiędzy atomami do odległości przyjętej jako kryterium odcięcia. Należy także uwzględnić oddziaływania elektrostatyczne, które dają istotny wkład przy niewielkich odległościach pomiędzy atomami. Dlatego tworzy się kompletne opisy dla aminokwasów i jednostek peptydowych NH-CO zawierające ładunki cząstkowe dla wszystkich atomów, ładunki wynikające z protonowania lub dysocjacji oraz wszystkie stałe siłowe dla deformacji geometrii. Stałe siłowe mogą być uzyskane eksperymentalnie (np. w oparciu o dane spektroskopowe) lub z precyzyjnych obliczeń teoretycznych. Takie zbiory parametrów opisujących składniki układu nazywa się polami siłowymi. Pola siłowe najczęściej stosowane do badań białek i kwasów nukleinowych to AMBER i CHARMM, ale znanych jest wiele innych. Minimalizacja energii układu nie wymaga bardzo długich obliczeń – pole siłowe pozwala na rozładowanie lokalnych naprężeń przez zmianę konformacji i niewielkie deformacje geometrii walencyjnej. Należy zauważyć, że pola siłowe muszą zachować podstawy fizyczne, w tym proporcje między stałymi siłowymi dla parametrów twardych (wiązania, kąty walencyjne) i miękkich (kąty torsyjne i oddziaływania niewiążące) gdyż te ostatnie powinny się zmieniać łatwiej. W przypadku kompleksu kilku cząsteczek (np. białko-ligand) parametryzacja wszystkich składników musi być spójna. W przeciwnym przypadku ‘sztywna’ cząsteczka liganda mogłaby przejść przez ‘miękką’ cząsteczkę białka jak nóż przez masło, co w oczywisty sposób byłoby niefizyczne i dało nieprawdziwe wyniki symulacji. Modelowanie metodą MM można zastosować, jeśli układ jest bardzo podobny do szablonu strukturalnego. To

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

ograniczenie wyniku z problemu, czy uda się osiągnąć globalne minimum energii układu wprowadzając niewielkie zmiany geometrii (głównie konformacyjne).

Takie ograniczenie podpowiada konieczność wzbogacenia arsenału narzędzi o dodatkowy składnik – dynamikę cząsteczki uwzględniającą energię kinetyczną atomów. Ta energia wynika z dwóch przyczyn – temperatury układu oraz wektora prędkości wynikającego z oddziaływań odpychających. Uwzględnienie tego członu w wyrażeniu na energię całkowitą układu prowadzi do metody Dynamiki Molekularnej MD. W tej metodzie obok pól siłowych opisujących energetykę oddziaływań między atomami i między ładunkami stosuje się równania ruchu Newtona. W przypadku kolizji między fragmentami cząsteczki (najczęściej łańcuchy boczne) energia kinetyczna wynikająca z odpychających oddziaływań może rosnąć do wysokich wartości. Zgodnie z termodynamiką statystyczną można to przeliczyć na temperaturę układu, która w skrajnych przypadkach symulacji może rosnąć do nieskończoności. Dlatego w przypadku symulacji MD często skaluje się prędkości zakładając stałą temperaturę, np. 300K. Nawet pobieżna analiza porównawcza wskazuje, że metoda MD nadaje się do minimalizacji energii układów dalszych od rzeczywistej struktury niż te możliwe dla metody MM. Można ten proces opisać jako dostarczenie energii kinetycznej niezbędnej do wyjścia z lokalnego minimum energii, czyli pokonania bariery potencjału, i osiągnięcia minimum globalnego. Wariantem metody MD jest procedura Simulated Annealing SA. Polega ona na dostarczeniu układowi energii kinetycznej przez podwyższenie temperatury, symulację MD w tej temperaturze ze skalowaniem prędkości, po czym stopniowe powolne schładzanie układu, aby po przekroczeniu lokalnych barier potencjału układ osiągnął minimum globalne bez wyjścia z odpowiadającej mu studni potencjału. Metoda MD-SA ma największy promień zbieżności, czyli może doprowadzić do prawdziwej struktury przestrzennej z najbardziej zdeformowanego modelu, w porównaniu z metodami MD i najsłabszą MM. Należy też podkreślić, że w optymalizacji geometrii cząsteczki białka tymi metodami trzeba uwzględnić otoczkę hydratacyjną, w tym technikę Periodic Boundry Conditions (PBC) pozwalającą zachować stałą masę układu. Wypada też wspomnieć o stosowaniu metod chemii teoretycznej, w szczególności DFT do optymalizacji struktur biomakrocząsteczek.

Uzyskane modele, tak jak i struktury określone na podstawie eksperymentu (krystalografia, NMR) poddaje się weryfikacji stosując kryteria stereochemiczne oraz fizyczne. Weryfikację poprawności geometrii cząsteczki można wykonać stosując popularny program PROCHECK. W przypadku modelowania cząsteczkowego MM i MD należy sprawdzić lokalne otoczenie każdego aminokwasu (hydrofobowe, polarne) stosując jeden z wielu programów, np. PROSA. W każdym przypadku można zweryfikować poprawność sfałdowania białka sprawdzając konformację łańcucha głównego przez analizę wykresu Ramachandrana opisującego  $\phi$  i  $\psi$  dla białka.

## Wykład 7/8: Zastosowanie metod analizy strukturalnej

Na koniec cyklu wykładów przeprowadźmy podsumowanie metod proteomiki strukturalnej. Proteomika strukturalna to część badań proteomiki poświęcona określeniu przestrzennej struktury

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

białek oraz użyciu tej informacji między innymi do poznania molekularnych podstaw aktywności enzymów, mechanizmów reakcji enzymatycznych, a także do projektowania małowcząsteczkowych związków regulujących aktywność biomakrocząsteczek (potencjalne leki). Tym celom służą metody dyfrakcyjne (krystalografia rentgenowska, elektronografia, neutronografia) metody modelowania molekularnego, magnetycznego rezonansu jądrowego (NMR) i mikroskopii elektronowej (EM), oraz modelowanie molekularne.

Pierwszym etapem badań jest najczęściej zastosowanie bioinformatyki do identyfikacji nieznanego białka albo do analizy podobieństwa do biomakrocząsteczek o znanej sekwencji, aktywności, funkcji lub strukturze. Zastosowanie globalnego lub lokalnego porównania sekwencji oparte na macierzach podobieństw mogą prowadzić do różnych wyników, dlatego wyboru metody postępowania trzeba dokonać starannie, w zależności od stawianych pytań.

W badaniach strukturalnych metodą z wyboru powinna być krystalografia rentgenowska. Z racji użyteczności i przeważającej liczby danych strukturalnych uzyskanych z zastosowaniem dyfrakcji promieniowania rentgenowskiego, tej metodzie poświęcono dużą część wykładu. Otrzymanie monokryształów białka (3D, czasem 2D) otwiera możliwość zastosowania metod dyfrakcyjnych. Otrzymanie periodycznej sieci krystalicznej tworzonej przez obiekty wykazujące wewnętrzną dynamikę (białka, kwasy nukleinowe) jest trudne. Po pierwsze trzeba makrocząsteczki utrzymać w roztworze. Po drugie należy doprowadzić do nasycenia roztworu a następnie do jego powolnego przesylenia i utworzenia nielicznych zarodków krystalizacji. W tym celu trzeba kontrolować przesylenie i szybkość wzrostu kryształu. Powolny wzrost w warunkach bardzo bliskich równowadze ciało stałe - roztwór, stwarza możliwość przyjęcia przez każdą cząsteczkę konformacji odpowiadającej minimum energii układu i utworzenia periodycznej sieci krystalicznej. Szybkie przekroczenie granicy nukleacji spowoduje utworzenie dużej liczby niewielkich kryształów i prawdopodobnie słabe uporządkowanie w ziarnach krystalicznych, co uniemożliwi określenie struktury kryształu.

Promieniowanie rentgenowskie oddziałuje z elektronami atomów tworzących cząsteczkę. Dlatego dane dyfrakcyjne pozwalają na obliczenie mapy rozkładu gęstości elektronowej. Ze względu na niewielką wartość atomowego czynnika rozpraszania atomów wodoru, ich wkład do rozpraszania jest niewielki. Dlatego lokalizacja atomów wodoru zwykle jest niemożliwa. Przeciętna rozdzielczość danych dyfrakcyjnych to 2.5-1.8 Å. Jednak użycie dużych monokryształów białka w eksperymencie z zastosowaniem synchrotronowych źródeł promieniowania pozwala uzyskać dane dyfrakcyjne o rozdzielczości atomowej (0.9-0.8 Å). Transformata Fouriera obliczona dla takich danych pozwala lokalizować atomy wodoru na mapach różnicowych. Znaczenia lokalizacji atomów wodoru nie da się przecenić. W większości reakcji mamy do czynienia z transferem atomu wodoru (jony wodorowe lub wodorkowe), udziałem grup zawierających protony w reakcjach enzymatycznych. Określenie położenia atomów wodoru pozwala więc poznać mechanizm reakcji katalizowanych przez biomakrocząsteczki, a następnie



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

użyć tej wiedzy dla projektowania skutecznych leków. Rozdzielczość struktur określonych metodą dyfrakcji promieniowania rentgenowskiego zawiera się w zakresie 10 – 0.7 Å. Wielkość obiektu badań jest także sporym wyzwaniem. W większości struktur zbadanych metodą krystalografii rentgenowskiej część asymetryczna struktury zawiera do 100 kDa białka lub kwasów nukleinowych. Największe struktury zawierają nawet 2-5 MDa. Ograniczeniem jest rozpuszczalność badanego białka. Szczególnie trudne jest otrzymanie kryształów białek membranowych, ze względu na ich dużą hydrofobowość powodującą asocjację i w konsekwencji wytrącanie. Dlatego do krystalizacji białek membranowych stosuje się roztwory zawierające PEG i detergenty.

W banku PDB jest zdeponowanych 40 struktur określonych metodą dyfrakcji elektronów. Dyfrakcja wiązki elektronów na monokryształach (elektronografia) pozwala uzyskać dane o niższej rozdzielczości 10-2 Å, choć zwykle do około 5-7 Å. Dlatego jest zrozumiałe, że takie dane mają ograniczone zastosowanie do projektowania leków. Użycie wiązki elektronów ma jednak ważną zaletę. Wiązka elektronów oddziałuje z elektronami atomów badanego obiektu. Oznacza to szczególne oddziaływanie z grupami naładowanymi (ujemnie lub dodatnio). Dyfrakcja elektronów pozwala więc zlokalizować grupy karboksylanowe czy amoniowe, a te mogą być istotne dla mechanizmu reakcji enzymatycznych, transmisji sygnałów (kinazy, kanały jonowe itp). Krystalografia elektronowa jest więc istotnym narzędziem proteomiki strukturalnej. Wielkość obiektu badań powinna być podobna jak dla krystalografii rentgenowskiej.

Bank PDB zawiera 44 struktury określone metodą dyfrakcji neutronów. Neutrony oddziałują z jądrami atomowymi. Dlatego w eksperymentach neutronograficznych atomy wodoru są ‘ciężkie’ tak jak atomy węgla, tlenu czy siarki. To pozwala na ich precyzyjną lokalizację, co jest ważne np. dla poznania mechanizmów reakcji enzymatycznych. Zdeponowane w PDB struktury mają rozdzielczość 3-1.5 Å, a jedna ma rozdzielczość 1.05 Å. Ze względu na zniszczenia kryształu wywołane wiązką neutronów, ta metoda nie jest często stosowana. Ograniczenia wielkości obiektu badań są podobne do tych dla krystalografii rentgenowskiej.

Zastosowanie krystalografii wymaga otrzymania dobrej jakości monokryształów o uporządkowanej sieci krystalicznej, mogących stanowić siatkę dyfrakcyjną. Często jest to bardzo trudne lub niemożliwe. Wtedy można zastosować mikroskopię elektronową. Dodatkową zaletą mikroskopii elektronowej jest fakt, że pozwala ona badać obiekty znacznie większe. W tej metodzie kieruje się wiązkę elektronów na badany obiekt. Preparat przygotowuje się przez naniesienie na płytkę roztworu lub zawiesiny białka, uzyskując kryształy dwuwymiarowe lub rozproszone pojedyncze cząsteczki. Dla kryształów można uzyskać obraz dyfrakcyjny i zastosować metody krystalografii elektronowej. W przypadku pojedynczych cząsteczek dokonuje się obrazowania przez transformację Fouriera. W tym przypadku można powiedzieć, że im większa cząsteczka, tym łatwiej uzyskać obraz. Wynikiem takich eksperymentów jest określenie kształtu cząsteczki, a w procesie rekonstrukcji trójwymiarowej z wielu obrazów przypadkowo położonych cząsteczek można uzyskać informacje o przestrzennej budowie cząsteczki. Rozdzielczość tej metody jest zwykle

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

ograniczona do 25-10 Å, choć w precyzyjnych eksperymentach uzyskano rozdzielczość dochodzącą do 4-5 Å. Z oczywistych względów taka informacja strukturalna nie może służyć projektowaniu leków. Jednak zastosowanie struktur fragmentów (podjednostek) badanego białka określonych z wysoką rozdzielczością metodami krystalografii rentgenowskiej i dopasowanie ich do obliczonego kształtu badanej cząsteczki pozwala na uzyskanie modelu pseudo-atomowego. W banku PDB jest ponad 700 struktur określonych metodą mikroskopii elektronowej. Tu zachodzi możliwość połączenia EM z krystalografią rentgenowską. Można wykorzystać taki model do obliczenia faz refleksów o niskiej rozdzielczości w krystalografii rentgenowskiej dla monokryształów badanego białka. Procedura rozszerzenia zbioru faz w dalszym etapie pozwala na uzyskanie struktury krystalograficznej dużych obiektów (np. 1 MDa) z rozdzielczością 3-5 Å. Tu znowu ograniczeniem połączenia EM z krystalografią jest możliwość otrzymania kryształów dużych białek.

Jeśli obiekt badań nie krystalizuje i jest za mały dla mikroskopii elektronowej, można zastosować Magnetyczny Rezonans Jądrowy NMR. Jądra atomów o niezerowym spinie dają w spektroskopii NMR sygnał, a najczęściej stosowana jest spektroskopia  $^1\text{H}$  NMR. Stosując zewnętrzne pole magnetyczne i wzbudzenie promieniowaniem radiowym można spowodować rozszczepienie jądrowych poziomów spinowych i przejścia między nimi. Występuje efekt ekranowania powodowany przez otoczenie danego jądra atomowego. Dlatego różne grupy dają sygnał w różnych miejscach widma. Dodatkowo mogą zachodzić sprzężenia pomiędzy blisko położonymi jądrami powodujące powstanie na widmie multipletów. Występuje też efekt NOE związany ze wzajemnym oddziaływaniem jąder ‘przez przestrzeń’, co pozwala powiązać położenie pasm z otoczeniem przestrzennym, a więc z konformacją cząsteczek.

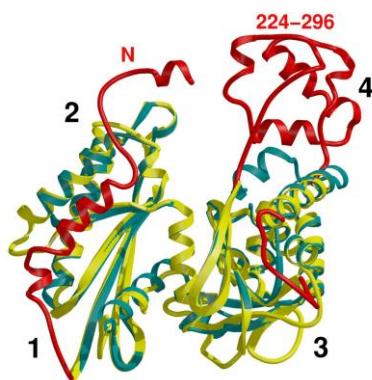
Widmo dla związków małowczątkowych jest stosunkowo proste, Jednak dla białek widmo jest bardzo złożone, co wynika z liczby atomów mogących dać sygnał na widmie NMR oraz ze zmienności otoczenia tych atomów. W spektroskopii NMR białek stosuje się kilka metod uproszczenia widma. Można prowadzić syntezę białka wykorzystując deuterowane aminokwasy. Deuter daje słabe sygnały w widmie NMR i to przy innych częstościach wzbudzenia niż Prot  $^1\text{H}$ . Stąd pomiar przy częstościach wzbudzających dla protu da widmo białka deuterowanego znacznie prostsze. Kolejną metodą uzyskania interpretowalnego widma jest zastosowanie odpowiedniej sekwencji impulsów radiowych, które mogą spowodować zniesienie rozszczepienia na multiplety, wygasić sygnał od niektórych grup lub wywołać przeniesienie namagnesowania na sąsiednie jądra. Techniki impulsowe połączone z deuteryzacją i połączeniem widm uzyskanych dla różnych jąder (np.  $^1\text{H}$ - $^{13}\text{C}$ ,  $^1\text{H}$ - $^{15}\text{N}$ ) pozwalają na zbudowanie przestrzennej struktury białka. Badania NMR białek prowadzi się w roztworach. Stąd ograniczenia wynikają z konieczności utrzymania białka w roztworze, co zwykle oznacza limit masy cząsteczkowej do 30-50 kDa. Największą strukturą badaną metodą NMR jest struktura kompleksu białka opiekuńczego i ko-chaperonu GroEL-GroES określona przez Kurta Wutricha, z masą cząsteczkową 900 kDa. W tych badaniach zastosowano częściową deuteryzację, techniki impulsowe, pomocna była symetria cząsteczki, co spowodowało uśrednienie sygnałów NMR.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

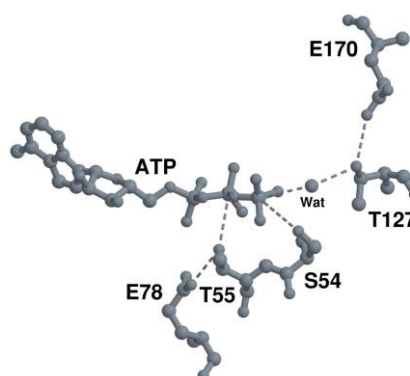
Jeśli nie możemy przeprowadzić eksperymentu ze względu na niemożność uzyskania czystego preparatu, pozostają techniki modelowania molekularnego. Modelowanie na podstawie homologii sekwencji może prowadzić do użytecznego modelu struktury, o ile identyczność sekwencji nie jest niższa od 30%. Jeśli stosuje się modelowanie ab initio, należy do uzyskanego modelu podchodzić bardzo ostrożnie. Wskazują na to choćby wyniki konkursów CASP (Critical Assessment of Techniques for Protein Structure Prediction) gdzie dla niektórych białek uzyskane modele są bardzo odległe od rzeczywistej struktury.

Wykład zakończymy prezentacją kilku przykładów wyników badań własnych z omawianej dziedziny.

Od kilkunastu lat, we współpracy z grupą biochemików prowadzone są badania nad roślinną NTPDazą, zwaną też apirazą. Jest to difosfohydrolaza ATP, hydrolizująca ATP i ADP do AMP i jonów ortofosforanowych. Białko to występuje w wielu organizmach, od bakterii, przez rośliny, aż do zwierząt. Wydaje się, że funkcją tego białka w organizmach jest kontrola poziomu di- i trifosfonukleotydów adeninowych. Te z kolei są neurotransmiterami, powodują perforację błony komórkowej, agregację płytek Kwi czy regulują poziom energii komórek. Dlatego aktywność NTPDaz łączy się z etiologią szeregu chorób. Waga problemu skłania do badań mechanizmu działania NTPDaz, określenia cząsteczkowych podstaw specyficzności substratowej oraz wpływu małowiązujących ligandów na aktywność enzymatyczną. Przez długi czas nie udało się otrzymać monokryształów NTPDazy<sup>1</sup>. Podjęto próby znalezienia strukturalnego szablonu do modelowania homologicznego. Po kilku latach udało się znaleźć nowo opublikowaną strukturę egzopolifosfatazy (PDB 1T6C), która jako jedyna wykazała istotną identyczność sekwencji z apirazą. Analiza podobieństw sekwencji ujawniła też, że NTPDaza należy do super-rodziny aktywny. Przeprowadzone modelowania pozwoliło z około 100 tys. modeli wybrać najlepszy. Model ten powstał przez zbudowanie nałożonych fragmentów na szablonie 1T6C, podobne modelowanie brakującej domeny na szablonie heksokinazy oraz zbudowaniu pętli łączących. Bardzo ważne były też wyniki przewidywania drugorzędowych elementów struktury, które okazały się spójne ze strukturą szablonów. Struktura została poddana minimalizacji metodą SA-MD. Tak uzyskany model został zweryfikowany z użyciem kryteriów stereochemicznych i poprawności sfałdowania dającego lokalne otoczenie każdego łańcucha bocznego.



Rys. 14.1



Rys. 14.2

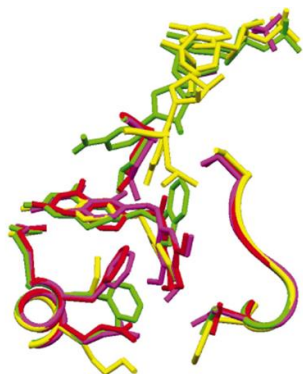
Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Ostateczny model zastosowano do dokowania substratu ATP i przeciwjonu wapniowego (**Rys. 14.1 i Rys. 14.2**). Modelowanie pozwoliło na wskazanie centrum katalitycznego enzymu oraz na identyfikację reszt aminokwasowych ważnych dla reakcji enzymatycznej. Na tej podstawie zaproponowano molekularne podstawy mechanizmu reakcji enzymatycznej NTPDazy1. Porównanie z danymi dotyczącymi aktywności w super-rodzinie aktywny pozwoliło zrozumieć podstawy specyficzności substratowej tych enzymów. Szczegóły można znaleźć w pracy [Kozakiewicz, A., et al. (2008) *ABP*, **55**, 141–150].

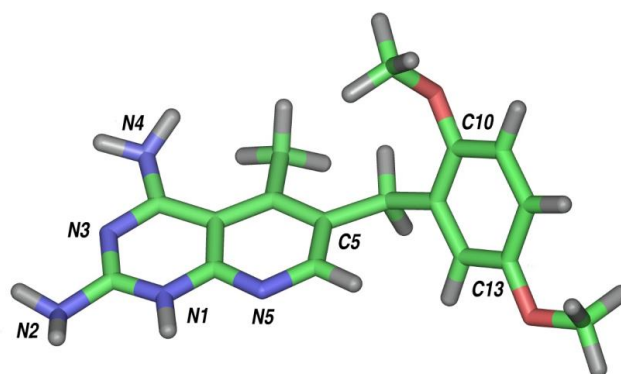
Kolejnym obiektem zainteresowań jest reduktaza dihydrofolianowa DHFR. Jest to enzym uczestniczący w cyklu transferu grupy metylowej, a to jest element syntezy puryn i aminokwasów komórkach. Wysoką aktywność DHFR wykazują szczególnie najszybciej rozwijające się komórki – nowotworowe lub bakteryjne w przypadku infekcji. Struktura krystaliczna DHFR z różnych organizmów jest znana. Porównanie struktur enzymu ludzkiego i bakteryjnych miało dać podstawy dla projektowania specyficznych inhibitorów enzymów bakteryjnych jako potencjalnych leków w przypadku infekcji. Znacznie trudniejszy przypadek stanowią nowotwory, gdyż enzym jest identyczny jak w zdrowych komórkach. Enzym DHFR jest elementem cyklu enzymatycznego przemian kwasu foliowego, w którym biorą udział także inne enzymy wiążące kwas foliowy. Obiektem zainteresowania były w szczególności synteza tymidylanowa TS oraz FPGS, enzym powodujący poliglutamylację kwasu foliowego. Taka przemiana jest niezbędna do transportu i rozpoznania folianów przez enzymy omawianego cyklu. Pierwotnym pomysłem było opracowanie inhibitorów blokujących dwa lub trzy wspomniane enzymy i w konsekwencji cały cykl enzymatyczny, co prowadziłoby do śmierci komórek lub blokowało ich rozwój. Otrzymano inhibitory nie podlegające poliglutamylacji (blokowanie FPGS oraz DHFR) albo posiadające podstawnik w centralnej części cząsteczki (PABA) uniemożliwiający bliskie położenie koenzymu NADPH, niezbędne to transferu jonu wodorkowego i redukcję DHF to tetrahydrofolianu THF. Badania strukturalne kilku kompleksów ludzkiej DGHR z takimi inhibitorami pokazały poprawność przyjętych założeń. W terapii przeciwnowotworowej najczęściej stosowanym inhibitorem jest metotreksat MTX. Jest to lek o dużej cytotoksyczności, a więc powodujący wiele skutków ubocznych. Okazało się, że komórki nowotworowe wykształcają lekooporność na MTX. Badania wykazały, że występuje w nich szereg mutacji w centrum katalitycznym związanych z podstawieniem Leu22, Glu30 i Phe31. W szczególności zamiana Glu30 na inny aminokwas znosi możliwość utworzenia podwójnego mostka solnego między MTX i DHFR, co powoduje obniżenie stałej wiązania metotreksatu. W rezultacie, MTX stawał się trucizną dla chorego, a nie komórek nowotworowych. Występujące mutacje Phe31 wiązały się z podstawieniem na małe aminokwasy: Gly, Ala, Ser. Badania strukturalne pozwoliły zrozumieć mechanizm tej lekooporności. Łańcuch boczny Phe31 oddziałuje z centralnym fragmentem PABA dihydrofolianów lub ich analogów (MTX) [Cody, V., Wojtczak, A., Kalman, T.I., Freisheim, J.H. & Blakley, R.L., in: *Chemistry and Biology of Pteridines and Folates*, Ayling, J.E., Nair, M.G. & Baugh, C.M. (eds), Plenum Press, pp 481-486 (1993).]. Wprowadzenie mniejszego łańcucha bocznego nie tylko likwidowało

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

oddziaływania  $\pi \dots \pi$  enzymu z MTX, ale stwarzało niszę w centrum katalitycznym (**Rys. 14.3**). W tej niszy wiązane były cząsteczki wody, co dodatkowo wpływało na obniżenie stałej wiązania. Jeden z inhibitorów pozwalał na pokonanie omawianej lekooporności. Był to piritrexim PTX (**Rys. 14.4**). Struktura jego kompleksu z hDHFR oraz wariantami enzymu została określona krystalograficznie.



**Rys. 14.3**



**Rys. 14.4**

Analiza strukturalna wykazała, że PTX tworzy alternatywne oddziaływania z resztami proliny w centrum katalitycznym, przez co jest efektywnie wiązany nawet w wariantach reduktazy dihydrofolianowej wykazujących oporność na metotreksat. We współpracy z grupą teoretyków przeprowadzone zostały obliczenia metodą dynamiki molekularnej dla uzyskania informacji o energetyce oddziaływań pomiędzy inhibitorem PTX i otaczającymi resztami. Wyniki pokazały, że istotne dla wiązania piritreksimu są nie tylko oddziaływania z resztami aminokwasowymi, ale także oddziaływania z koenzymem NADPH [Nowak et al. (2001), *ABP* **48**, 903–916]

Ostatnim przykładem są badania krystalograficzne transtyretyny TTR. Transtyretyna jest białkiem surowicy krwi. Jej funkcją jest transport hormonów tarczycy, zabezpieczenie małego białka RBP przed usunięciem przez nerki, a to przez utworzenie kompleksu TTR-RBP, oraz transport w krwiobiegu antybiotyków, flawonoidów i innych związków aktywnych biologicznie. Transtyretyna jest białkiem homotetramerycznym, z dwoma miejscami wiązania hormonów. W przypadku wystąpienia punktowych mutacji tetrametr ulega rozpadowi, a monomery w zmienionej konformacji asocjują tworząc włókna amyloidu, odkładające się w różnych tkankach. Taka amyloidoza może mieć podłoże genetyczne lub może występować spontanicznie (choroby starcze), jednak bez względu na przyczynę prowadzi do dysfunkcji organów i śmierci. Stwierdzono, że wiązanie hormonów tarczycy (T4, T3), ich analogów i kompetytorów stabilizuje tetrametr TTR, opóźniając wystąpienie amyloidozy. Jednak przy wiązaniu ligandów obserwuje się efekt ujemnej kooperatywności: wiązanie pierwszej cząsteczki liganda osłabia około 100-krotnie wiązanie w drugim miejscu. Szczegółowe badania strukturalne pozwoliły na określenie podstaw mechanizmu ujemnej kooperatywności

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

[Neumann, P., Cody, V. & Wojtczak, A. (2001). *Acta Biochim. Polonica*, **48**, 867-875.].  
Badania wskazały także, że jeden z analogów hormonu, kwas tetrajodotyreoctowy T4Ac, jest wiązany jednocześnie w obu miejscach wiązania,. W ten sposób T4Ac stabilizuje cały tetrametr TTR, bez względu na możliwość efektu ujemnej kooperatywności.

Wnioskiem końcowym z cyklu wykładów może być stwierdzenie, że najlepsze rezultaty daje połączenie kilku metod. Każda z nich ma zalety, ale ma też istotne ograniczenia. Użycie kilku może pozwolić na obejście tych ograniczeń i uzyskanie wyników istotnych dla rozwoju wiedzy, ale także o aspektach praktycznych.

## 9. Literatura do części I

1. Z. Bojarski [et al.]. *Krystalografia : podręcznik wspomagany komputerowo* Wyd. Naukowe PWN, 2001.
2. M. Jaskólski, *Krystalografia dla biologów*. Wydawnictwo Naukowe UAM, 2010.
3. P. Luger, *Rentgenografia strukturalna monokryształów*, PWN, 1989.
4. *International Tables for X-ray Crystallography Vol. A. Space group symmetry*. Kluwer, 1995.
5. *International Tables for X-ray Crystallography Vol. B. Reciprocal space*. Kluwer, 1993.
6. *International Tables for X-ray Crystallography Vol. F. Crystallography of biological macromolecules*. Kluwer, 2001.
7. Paul G. Higgs, Teresa K. Attwood, *Bioinformatyka i ewolucja molekularna*. PWN, 2011.
8. A. D. Baxevanis i B. F. F. Ouellette (red). *Bioinformatyka : podręcznik do analizy genów i białek*. PWN, 2005.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Część II

### Wykład 1: Projektowanie leków - krystalografia jako narzędzie w chemii, biochemii i medycynie

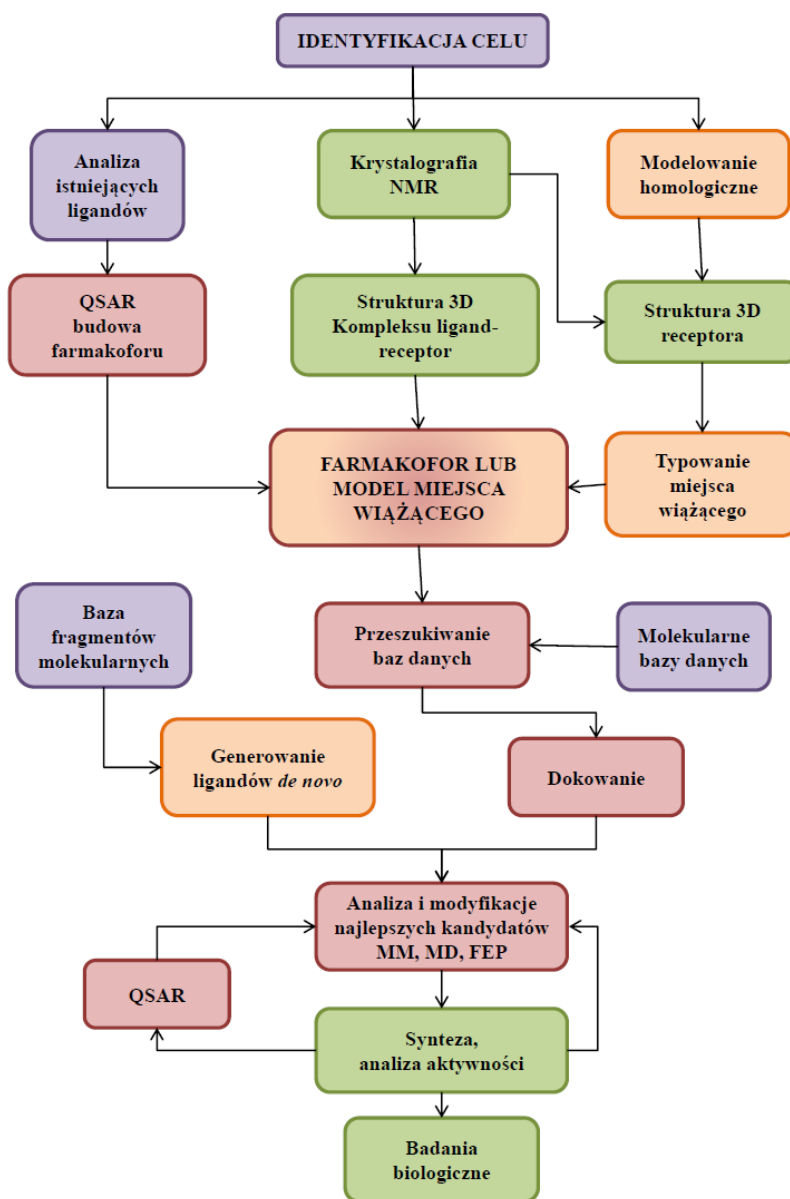
Ciągły wzrost mocy obliczeniowej komputerów, rozwój metod i narzędzi bioinformatycznych oraz rozbudowa strukturalnych baz danych małych molekuł (ligandów) oraz białek sprawiają, że projektowania leków z użyciem komputerów (Computer Aided Drug Design) staje się coraz efektywniejszym i coraz częściej stosowanym narzędziem. Obecnie stosowane metody, nie zawsze pozwalają na pewne i wiarygodne zaproponowanie ostatecznej struktury działających leków, umożliwiają natomiast wybór obiecujących kandydatów spośród milionów cząsteczek dostępnych w bazach danych oraz ich optymalizację pod kątem oddziaływania z białkiem. Można w ten sposób obniżyć koszt a zarazem czas potrzebny na zaprojektowanie nowego leku. Pierwszy etap procesu projektowania nowej cząsteczki chemicznej, która w przyszłości ma pełnić rolę leku jest dość trudny do określenia. Punktem startowym może być choroba i dokładne poznanie jej przyczyn, mechanizmu i skutków bądź teoretyczna weryfikacja już istniejących struktur i doświadczalne poszukiwanie zmodyfikowanych cząsteczek.

W zależności od wiedzy, którą posiadamy na wstępnym, startowym etapie procesu projektowania leków wyróżnić można dwie główne drogi:

- projektowanie w oparciu o strukturę znanych ligandów (ligand based design), na które składają się następujące etapy: budowa farmakoforu, analiza QSAR,
- projektowanie w oparciu o znaną strukturę receptora (receptor based design), w skład którego wchodzi analiza budowy miejsca wiążącego, analiza zmienności konformacyjnej receptora,

Te dwie ścieżki dotyczą tylko wstępnych etapów procesu projektowania, prowadzą do otrzymania modelu farmakoforu lub modelu miejsca wiążącego. Dalsze etapy takie jak: przeszukiwanie istniejących baz danych związków małowcząsteczkowych, dokowanie, otrzymywanie nowych ligandów przebiegają w analogiczny sposób niezależnie od startowych etapów. Rysunek 1.1 przedstawia różne drogi w procesie projektowania leku oraz metody eksperymentalne jak i obliczeniowe wykorzystywane w tym procesie. Jakość danych wejściowych w duży sposób wpływa na osiągnięcie sukcesu w procesie projektowania cząsteczki leku. Na wzrost ryzyka niepowodzenia może wpłynąć kilka etapów: modelowanie struktury receptora (typowanie miejsca aktywnego) lub generowanie nowych ligandów (rys. 1.1: etapy zaznaczone kolorem pomarańczowym). Potrzeba skonstruowania modelu receptora w przypadku braku jego struktury 3D lub nieznaną strukturę centrum aktywnego receptora wymagają zastosowania dodatkowych założeń i przybliżeń, co powoduje zwiększenie prawdopodobieństwa otrzymania błędnych wyników. Generując całkiem nowe ligandy częściej dostaje się struktury nie będące lekami, niż wtedy, gdy ogranicza się do modyfikacji znanego wcześniej związku wiodącego (lead compound).

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki



**Rysunek 1.1** Schemat wykorzystania metod eksperymentalnych i obliczeniowych w etapach projektowania leków.

Znajomość zarówno grupy ligandów danego receptora oraz ich aktywności jak i struktur 3D kompleksów białko-ligand oraz formy natywnej białka jest idealną sytuacją.

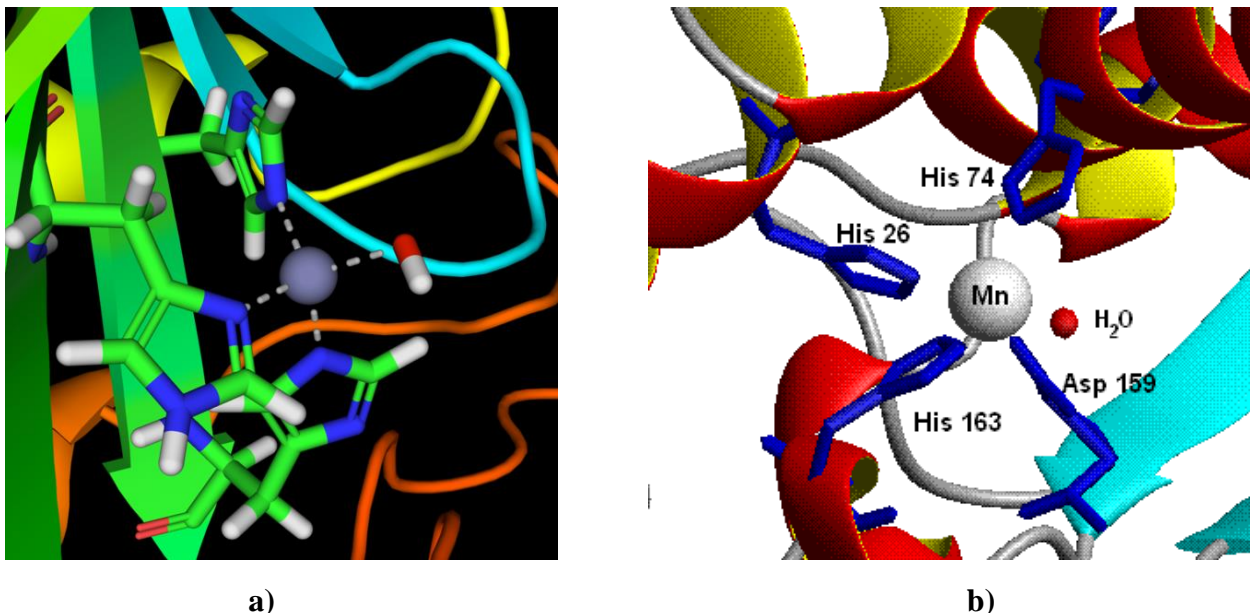
Leki projektowane z użyciem metod obliczeniowych istnieją już i są dostępne na rynku farmaceutycznym. Jako główny przykład można przytoczyć blokery receptorów dla angiotensyny – sartany.



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Typowanie miejsca wiążącego

Bardzo ważne znaczenie w projektowaniu cząsteczki leku odgrywa znajomość budowy centrum aktywnego receptora (Rysunek 1.2).



**Rysunek 1.2** a) Model miejsca aktywnego anhidrazy węglanowej. Widoczne trzy histydyny i jedna grupa hydroksylowa koordynujące jon cynku, będący kofaktorem tego enzymu. b) Centrum aktywne dysmutazy ponadtlenkowej (SOD).

Nie zawsze jednak dysponuje się strukturą 3D receptor - ligand, która dostarcza dokładnej informacji o położeniu i budowie miejsca wiążącego. Najczęściej poszukuje się centrum aktywnego w następujących przypadkach:

- znana jest tylko struktura krystalograficzna natywnej formy receptora,
- znana jest jedynie sekwencja aminokwasowa badanego białka,
- poszukuje się alternatywnego miejsca wiążącego,

W tych przypadkach można wykorzystać metodę opartą na homologii. Trzeba tylko uwzględnić, że nawet poprawne zaproponowanie miejsca wiążącego nie zapewnia sukcesu - czasami w wyniku związania liganda receptor ulega dużym zmianom konformacyjnym, które trudno przewidzieć bez wykonywania dokładnych symulacji.

## Metody oparte na homologii

W przypadku, gdy poszukuje się centrum aktywnego dla istniejących grupy ligandów danego białka można skorzystać z informacji, że będzie ono podobne do innych, opisanych już miejsc wiążących w białkach pokrewnych. W początkowym etapie należy porównać sekwencję badanego

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

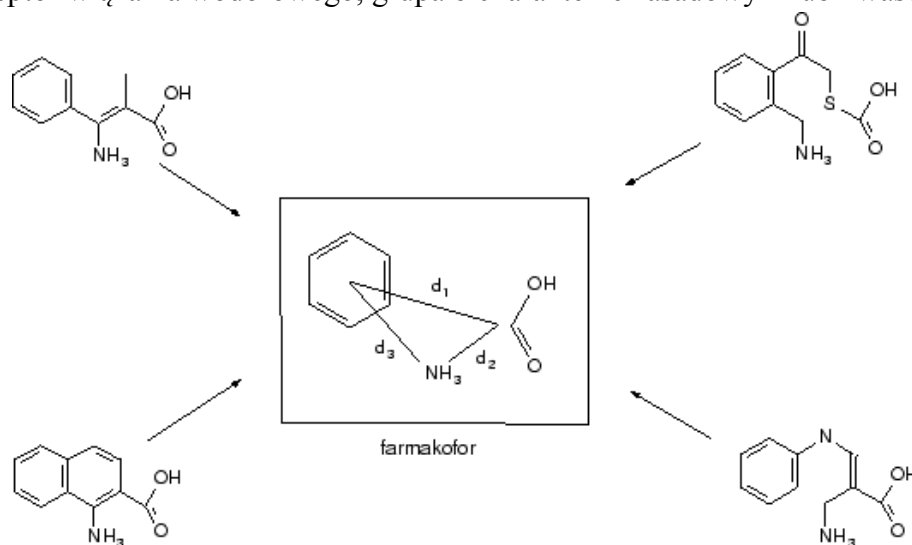
receptora z sekwencjami tych białek. W centrum aktywnym fragmenty aminokwasów występują zazwyczaj zachowane w procesie ewolucji, dlatego też znalezienie fragmentów o niewielkiej zmienności sekwencyjnej może sugerować, że to właśnie aminokwasy z tego fragmentu tworzą miejsce wiążące.

Gdy sekwencja badanego białka okaże się wyjątkowa, nieznana można jeszcze oczekiwać, że będzie ona strukturalnie podobna do białka z dobrze zdefiniowanym centrum aktywnym. W tym celu można wykorzystać bazę danych białek podobnych strukturalnie. Na przykład bazę FSSP (Families of structurally similar proteins), która jest dostępna bezpłatnie dla użytkowników akademickich. Do wyszukiwania białek podobnych strukturalnie można wykorzystać program Dali.

Projektowanie cząsteczki leku, które oparte jest na znajomości struktury 3D receptora (topologia centrum aktywnego) warto dodatkowo uzupełnić o informację pochodzące z analizy farmakoforu i metody QSAR.

## Budowanie farmakoforu

Farmakofor jest modelem zawierającym informacje o zależnościach przestrzennych między fragmentami wspólnymi dla ligandów oddziałujących z danym białkiem. Gdy znana jest struktura kilku aktywnych ligandów badanego receptora a nie dysponujemy informacjami o jego strukturze 3D etap tworzenie farmakoforu odgrywa bardzo istotną rolę w projektowaniu leków. W optymalnej sytuacji ligandy te powinny stanowić aktywną i różnorodną grupę związków. Najważniejszym warunkiem jest oddziaływanie badanych ligandów z tym samym centrum aktywnym receptora. Farmakofor składa się z fragmentów odgrywających ważną rolę w oddziaływaniach ligand-receptor (Rysunek 1.3). Mogą to być atomy szczególnych typów (np. atom azotu), grupy chemiczne (np. grupa hydroksylowa, aminowa, pierścień fenyłowy), lub też grupy o szczególnych funkcjach (np. donor lub akceptor wiązania wodorowego, grupa o charakterze zasadowym lub kwasowym, grupa).



Rysunek 1.3 Przykładowy schemat farmakoforu.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Zbudowany farmakofor może być użyty w wielu etapach procesu projektowania leków:

- do przeszukiwania baz danych strukturalnych,
- do tworzenia ligandów de novo,
- do optymalizacji molekuł wiodących,
- w metodzie 3D-QSAR,

Obecnie dostępna jest spora liczba programów komputerowych przy pomocy których, jeśli dysponuje się grupą ligandów danego receptora, można zbudować model farmakoforu. Opierając się na założeniu, że wszystkie molekuły oddziałują z tym samym centrum aktywnym, muszą generować ten sam farmakofor, tworzy się wspólny jeden model dla całej grupy. Hipoteza ta nie zawsze jednak się sprawdza. Mogą istnieć bardzo aktywne ligandy, które nie będą zawierały któregoś z fragmentów zbudowanego farmakoforu.

## Model QSAR - Deskryptory

QSAR (Quantitative Structure - Activity Relationships) jest to metoda umożliwiająca analizę zależności między aktywnością biologiczną grupy ligandów, a ich cechami fizykochemicznymi. Metoda ta pozwala zbadać, które właściwości tych ligandów mają największy wpływ na aktywność oraz sformułować matematyczny model tej zależności, zwany równaniem QSAR. Znajomość tego równania pozwala:

- zaproponować mechanizm oddziaływania ligandów z receptorem,
- wyliczyć aktywność niezbadanych jeszcze molekuł,
- zasugerować najlepsze z punktu widzenia aktywności zmiany w budowie ligandów,
- pogrupować ligandy pod względem ich aktywności oraz cech fizykochemicznych,

Wykonanie analizy statystycznej oraz zaproponowanie równania QSAR jest możliwe gdy wyrazimy parametry fizykochemiczne grupy badanych ligandów w postaci liczbowej, czyli w postaci tzw. deskryptorów.

Wiele różnych cech fizykochemicznych (rozmiar i ładunek drobiny, rodzaj grup funkcyjnych, rozpuszczalność) może wpływać na aktywność biologiczną związków. Przypisując tym cechą wartości liczbowe tworzymy funkcje zwane deskryptorami. Deskryptory, ze względu na charakter właściwości, które opisują można podzielić na kilka grup:

- właściwości strukturalne, np.: masa molowa, objętość molekularna, pole powierzchni, momenty bezwładności, liczba wiązań pojedynczych,
- właściwości elektronowe, np.: ładunek całkowity, moment dipolowy, polaryzowalność, energia HOMO, energia LUMO,
- właściwości termodynamiczne, np.: współczynnik podziału między wodę i oktanol, hydrofobowość podstawników, refraktywność molowa,

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

- właściwości przestrzenne pól - CoMFA (Comparative Molecular Field Analysis).

Analiza QSAR umożliwia wygenerowanie równania, które opisuje zależność aktywności biologicznej ligandów w funkcji poszczególnych deskryptorów.

### **Przeszukiwanie baz danych małych cząsteczek**

Przeszukując strukturalne bazy związków małowcząsteczkowych, z olbrzymiej grupy występujących w nich związków, można wybrać molekuly spełniające kryteria zapisane w modelu farmakoforu i modelu QSAR. W ten sposób można ograniczyć liczbę związków, które będą badane w kolejnych etapach procesu projektowania cząsteczki leku.

### **Dokowanie ligandów do receptora**

Dokowanie jest metodą, która pozwala przewidzieć sposoby ułożenia i konformację liganda w centrum aktywnym receptora oraz ocenić siłę oddziaływań międzycząsteczkowych. Zazwyczaj dla badanego receptora analizie poddaje się nawet do kilkunastu tysięcy ligandów, dlatego algorytmy wykorzystywane w procesie dokowania muszą być zarazem szybkie jak i dokładne.

Do przeprowadzenia dokowania potrzebna jest znajomość struktury receptora, którą można uzyskać na drodze analizy strukturalnej, widm NMR lub modelowania homologicznego. W procesie dokowania potrzebna jest także grupa potencjalnych ligandów. Niezbędna jest również znajomość topologii miejsca wiążącego, ponieważ dokowanie do całego receptora jest pozbawione sensu.

Centrum aktywne zazwyczaj zlokalizowane jest w kieszeni, szczeliny białka, tworzą je aminokwasy, które w sekwencji aminokwasowej mogą być znacznie oddalone od siebie, natomiast w cząsteczce trójwymiarowej znajdują się blisko siebie. Reszty aminokwasowe w miejscu wiążącym receptora są ewolucyjnie zachowane, ponieważ zmiana któregoś z tych aminokwasów mogłaby uniemożliwić enzymowi katalizowanie dotychczasowych reakcji.

Przed rozpoczęciem procedury dokowania należy dokładnie przeanalizować budowę miejsca wiążącego. Powinno się zwrócić uwagę na grupy donorowe i akceptorowe dla wiązań wodorowych. Dwudodatnie jony metali często występują w centrach wiążących receptorów, jeśli zostawimy je podczas dokowania, musimy zadbać o ich poprawną parametryzację (przypisanie właściwego ładunku). Cząsteczki wody pełnią istotną rolę w miejscu wiążącym, mogą brać udział w „mostkowaniu” wiązań wodorowych, dlatego warto je uwzględnić jako element centrum aktywnego w trakcie dokowania. Jeśli znamy aminokwasy, które mogą tworzyć miejsce wiążące warto jest zdefiniować więzy farmakoforyczne. Pozwalają one wskazać miejsca, w którym powinien zostać umieszczony określony fragment molekuly (np. donor wiązania wodorowego). Ligandy używane do dokowania powinny być również odpowiednio przygotowane, potrzebujemy pełnoatomową strukturę oraz atomy powinny mieć przypisany właściwy ładunek.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Projektowanie nowych ligandów i modyfikacja istniejących

Omawiane do tej pory metody, takie jak przeszukiwanie strukturalnych baz danych, tworzenie równania QSAR, bądź dokowanie, opierają się na analizie znanych ligandów w celu znalezienia w tej grupie najlepszych kandydatów na "molekuły wiodące" (lead compounds). Następnie związki te po serii udoskonaleń mają szansę stać się prototypem leku. Projektowanie w oparciu o strukturę znanych ligandów wymaga oczywiście posiadania bazy takich związków, co nie zawsze jest możliwe do zapewnienia.

Tworzenie nowych ligandów (de novo drug design) jest alternatywnym rozwiązaniem bazując na informacjach uzyskanych z analizy farmakoforu lub miejsca wiążącego. Niestety, istnieje około  $10^{80}$  możliwych do zsyntezowania związków, które mogłyby stać się lekiem. Każda z tych molekuł może przyjmować różne konformacje i na wiele sposobów układać się w centrum aktywnym receptora. Tyle zmiennych w trakcie generowania nowych ligandów może spowodować, że proces projektowania nowego leku nie zakończy się sukcesem. Projektowanie nowych ligandów polega na losowych zmianach struktury i położenia nowego liganda, a następnie ocenie jego ułożenia w zdefiniowanym wcześniej centrum aktywnym. Lepsze rezultaty można uzyskać modyfikując znane wcześniej molekuły wiodące niż tworząc nowe cząsteczki.

W procesie projektowania leków warto również pamiętać o prawie piątek Lipińskiego, zgodnie z którym leki doustne powinny spełniać następujące kryteria:

- co najwyżej 5 donorów wiązań wodorowych (np. grupy -NH, -OH),
- nie więcej niż 10 grup akceptorowych wiązania wodorowego (atom O lub N),
- masa cząsteczkowa związku mniejsza niż 500 Da,
- współczynnik podziału (n-oktanol-woda) mniejszy niż 5.

## Wykład 2: Otrzymywanie i oczyszczanie białek na potrzeby krystalografii.

Znane są trzy sposoby uzyskania białek do badań krystalograficznych: izolacja z oryginalnego materiału biologicznego, synteza chemiczna oraz białka uzyskane na drodze inżynierii genetycznej – białka rekombinowane.

### 1. Izolacja białka z oryginalnego materiału biologicznego

Rodzaj materiału biologicznego, metoda izolacji białka oraz sposób jego oczyszczania w dużej mierze zależy od celu jego zastosowania. W badaniach enzymatycznych wyizolowane białko powinno zachować swoją natywną formę oraz aktywność biologiczną. Badania strukturalne lub farmaceutyczne wymagają preparatu homogenicznie czystego. Należy wybrać materiał biologiczny zawierający jak największą ilość pożądanego białka, który w łatwy i tani sposób można otrzymać w roztworze w formie aktywnej. W otrzymanym roztworze białka powinno być jak najmniej zanieczyszczeń.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Obecnie obserwuje się duży rozwój technik i metod rozdzielania oraz oczyszczania białek, ma to ogromny wpływ na różne dziedziny nauki oraz przemysłu. Stopień czystości otrzymanego preparatu jest różny i zależy od jego potencjalnego wykorzystania. Badania krystalograficzne i terapeutyczne wymagają białek o wysokim stopniu czystości (95-99%), natomiast preparat białkowy stosowany w przemyśle spożywczym może cechować się mniejszym niż 95% stopniem czystości.

Praca z preparatem białkowym wymaga większej dokładności i delikatności niż praca ze związkami małocząsteczkowymi. Inny sposób izolacji białek związany jest przede wszystkim z ich właściwościami:

- białka są dużymi, elastycznymi i labilnymi związkami organicznymi, ich struktura trzeciorzędowa może być łatwo zniszczona w procesie denaturacji, który zarazem pozbawia je aktywności biologicznej. W związku z tym, w procesie izolacji oraz oczyszczania białka powinny być stosowane jedynie bardzo łagodne procedury. Używanie wysokich temperatur (destylacja, gotowanie) jest zabronione.
- białka zbudowane są z tych samych bloków budulcowych (aminokwasów) ułożonych jedynie w innej sekwencji, dlatego są one do siebie bardzo podobne. Zatem, stosowane techniki muszą być bardzo selektywne i subtelne.

## Trudności w izolacji białka

Produkcja białek zachodzi jedynie w organizmach żywych. Wyizolowanie białka wymaga jego uwolnienia z materiału biologicznego a w kolejnym kroku należy je oddzielić od pozostałych białek, co jest bardzo trudnym etapem. Wyizolowanie całego białka (bez strat) z wybranego materiału biologicznego oraz jego homogeniczne oczyszczenie jest optymalną sytuacją, co nie jest łatwe do wykonania. Procedurę izolacji białka należy starannie zaplanować, powinna ona obejmować jak najmniej etapów przeprowadzonych w jak najkrótszym czasie.

## Wybór źródła białka

Wyboru źródła białka dokonuje się po określeniu białka badanego, docelowego. Tkanki zwierzęce, roślinne, drożdże lub bakterie mogą być materiałem startowy. Zawartość białka oraz jego rozmieszczenie w znaczący sposób wpływa na wybór materiału biologicznego. Uwzględniając umiejscowienie docelowego związku, białka dzielimy na trzy grupy:

- białka znajdujące się w środku komórki:
  - w cytozolu,
  - w organellach komórkowych,
- białka występujące w płynach ustrojowych - białka rozpuszczalne,
- białka połączone z fragmentami stałymi komórki lub białka nierozpuszczalne:
  - białka strukturalne,
  - białka błonowe.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
 realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Izolowanie i ekstrakcja białka z wybranego materiału biologicznego

Uwolnienie docelowego białka do roztworu wymaga dobrania odpowiedniej procedury w zależności od jego umiejscowienia. Jeżeli białko związane jest z błoną komórkową lub występuje w organellach komórkowych należy zastosować dodatkowe etapy, które umożliwią rozbicie tkanek, jak i komórek. Do rozbijania komórek znajdujących się w zawieszynie można zastosować różne metody (Tabela 2.1).

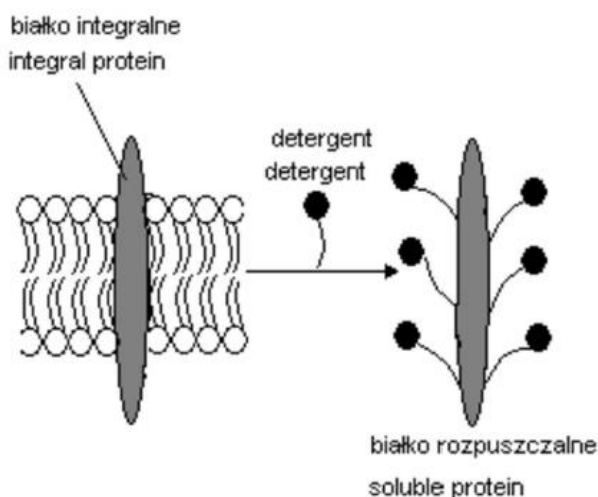
Do dyspozycji mamy łagodne, umiarkowane oraz energiczne metody. Najczęściej stosowanym metodami jest sonifikacja, homogenizacja, równie często używane jest wysokie ciśnienie. Ciekawą metodą rozbijania komórek jest tzw. „bomba azotowa”. W metodzie tej wykorzystuje się gwałtowną dekompresję, używając wysokiego ciśnienia, nasyconej rozpuszczalnym gazem (np. azotem) zawiesziny komórek. Uwalniając preparat do roztworu można również wykorzystać ogrzewanie, działanie detergentami lub rozpuszczalnikami organicznymi. W innej metodzie, wartej wspomnienia, stosuje się dyfuzję wody z roztworu izotonicznego do roztworu o większym stężeniu – cytozolu, co powoduje pęcznienie i rozrywanie komórek.

**Tabela 2.1** Metody uwalniania białek z materiału biologicznego do roztworu.

Metody uwalniania białek Methods of extracting and introducing proteins	Źródło białka Source of protein
Łagodna Gentle method	
Liza komórki (szok osmotyczny) Cell lysis (osmotic shock)	Erytrocyty, <i>E.coli</i> , okołoplazmowe: białka wewnątrzkomórkowe Erythrocytes, <i>E.coli</i> , periplasm: intracellular proteins
Trawienie enzymatyczne Enzymatic digestion	Bakterie: białka wewnątrzkomórkowe Bacteria: intracellular proteins
Ręczna homogenizacja Manual homogenisation	Tkanka wątroby Liver tissue
Siekanie (mielenie) Mincing (grinding)	Mięśnie Muscles
Umiarkowana Moderate method	
Homogenizator Blade'a Blade's homogeniser	Tkanka mięśniowa, większość tkanek roślinnych i zwierzęcych Muscle tissue, the majority of animal and plant tissues
Mielenie ze ścieraniem, np. piaskiem Grinding using abrasive materials, e.g. sand	Bakterie, tkanki roślinne Bacteria, plant tissues
Energiczna Vigorous method	
Ultradźwięki Ultrasonication	Zawiesina komórek Cell suspension
Homogenizator Manton-Gaulina Manton-Gaulin homogeniser	Zawiesina komórek Cell suspension
Prasa Frencha French press	Bakterie, komórki roślinne Bacteria, plant cells
Precypitacja frakcyjna Fractional precipitation	Lizat komórek, przeciwciała monoklonalne, białka zewnątrzkomórkowe Cells lysate, monoclonal antibodies, extra-cellular proteins

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

W procesie uwalniania do roztworu badanego białka błonowego powinien zostać zastosowany dodatkowy etap (Rysunek 2.1).



**Rysunek 2.1** Mechanizm uwalniania białek błonowych za pomocą detergentów.

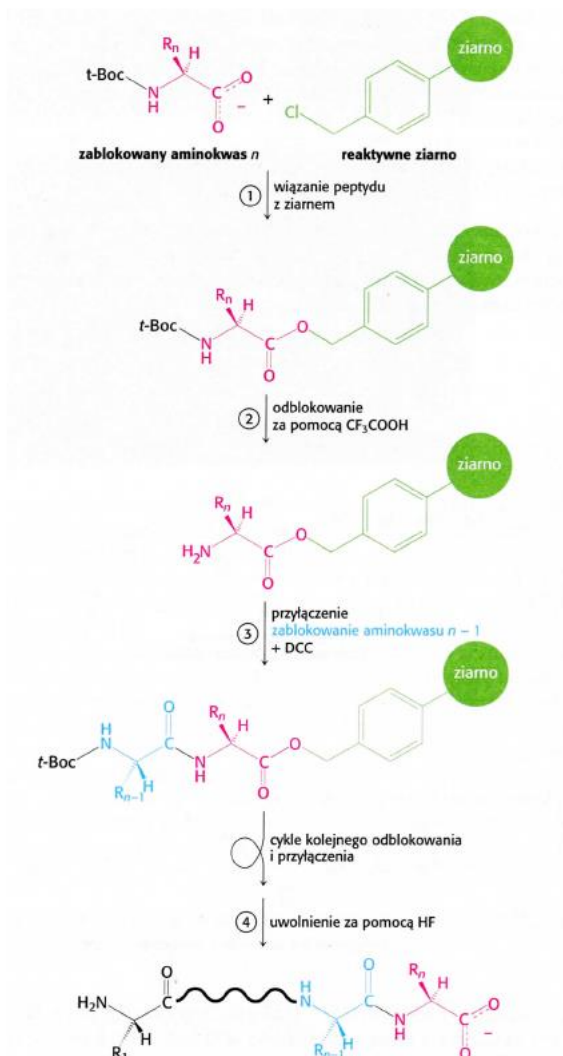
Zaledwie 1% białek błonowych ma poznaną strukturę, mimo, że stanowią one sporą liczbę wszystkich białek. Rozwój badań strukturalnych umożliwi powiększyć grupę znanych struktur tych białek. Jak wspomniano wcześniej, badania krystalograficzne wymagają preparatu o wysokim stopniu czystości. Zastosowanie detergentu anionowego powoduje uszkodzenie warstwy lipidowej błony komórkowej i uwolnienie białek integralnych do roztworu.

## 2. Synteza chemiczna

Białka, jak wiadomo, są liniowymi polimerami złożonymi z aminokwasów, co umożliwia otrzymanie ich na drodze syntezy chemicznej. Można by zatem łączyć, w reakcji kondensacji, aminokwasy zgodnie z sekwencją, co prowadziłoby do otrzymania odpowiedniego białka. Podczas syntezy stykamy się jednak z pewnym problemem. Reakcja ta nie zachodzi ze 100% wydajnością, ponieważ „cegiełki” budujące białko posiadają dodatkowe grupy funkcyjne, które również mogą brać udział w reakcji kondensacji. Jeśli chcielibyśmy otrzymać dipeptyd AB, to karboksylowa grupa cząsteczki A powinna przereagować z aminową grupą aminokwasu B. Jednak obie cząsteczki zawierają zarazem grupę karboksylową, jak i amonową, co skutkuje utworzeniem mieszaniny dipeptydów (pożądanego AB oraz AA, BB i BA) i nieprzereagowanych aminokwasów A i B. W ten sposób otrzymujemy trudną do rozdzielania mieszaninę z niewielką ilością docelowego produktu.



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki



**Rysunek 2.2** Synteza peptydu na podłożu stałym.

Problem ten został rozwiązany dzięki B. Merrifieldowi, który w 1962 roku zaproponował metodę syntezy na stałym podłożu (Rysunek 2.2.). Metoda ta wymaga zastosowania specjalnych grup blokujących, które uniemożliwiają tworzenie się niepożądanych produktów. Syntezę SPPS (Solid Phase Peptide Synthesis) rozpoczyna się od przywiązania grupą COOH ostatniego aminokwasu do stałego podłoża. Grupy aminowe w trakcie syntezy blokuje się grupą t-Boc (tert-butyloksykarbonylową) lub FMOC (9-fluorenylometoksykarbonylową). Po przyłączeniu aminokwasu do podłoża i przed dodaniem kolejnego w sekwencji aminokwasu, usuwana jest blokada z ostatniej reszty aminowej. Zmieniając pH roztworu możemy zdjąć blokadę z grupy aminowej, stosujemy środowisko kwaśne dla grupy t-Boc lub zasadowe dla FMOC. W kolejnym etapie, dokładnie odmywamy produkty niepożądane i nieprzereagowane substraty a interesujący nas peptyd, zostaje związany z podłożem. Cyklicznie powtarzając omówiony proces otrzymujemy interesujący nas produkt. Ostatnim etapem syntezy jest zdjęcie wszystkich blokad oraz odcięcie otrzymanego produktu od stałego podłoża przy użyciu TFA (kwasu trójfluorooctowego).

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Zaproponowana przez B. Merrifielda metoda wydaje się banalnie prosta ale zastosowanie jej w syntezie dużych białek nie pozwala jednak na otrzymanie produktu z dobrą wydajnością. Poza tym jest to bardzo pracochłonna i czasochłonna metoda. Obejmuje szereg procesów tworzenia wiązań peptydowych, zdejmowania grup blokujących oraz usuwania produktów ubocznych. Używając syntezy na stałym podłożu otrzymać możemy krótsze polipeptydy, które można następnie skleić, np. metodą ligacji chemicznej. Metoda ta odgrywa również istotną rolę w przypadku otrzymywania polipeptydów zbudowanych z niebiałkowych aminokwasów.

Proteaza retrowirusa HIV (ludzki wirus upośledzenia odporności) była pierwszym syntetycznym białkiem, dla którego struktura 3D została określona. Białko to jest stosunkowo nieduże (99 reszt) ale występuje w postaci homodimeru.

### 3. Białka rekombinowane

Dzięki rozwojowi inżynierii genetycznej możliwe jest stworzenie białek rekombinowanych, które znalazły szerokie zastosowanie w medycynie. Są one powszechnie używane w terapii wielu chorób onkologicznych, neurologicznych oraz autoimmunologicznych. Białka rekombinowane produkowane są w specjalnie przygotowanych komórkach. W ten sposób można tanio, szybko i bez większych problemów produkować duże ilości potrzebnego białka. Otrzymany w ten sposób preparat można w łatwy sposób oczyścić, pamiętając o dokładnym i indywidualnym zaplanowaniu procedury oczyszczania. Technika rekombinacji DNA odgrywa istotną rolę w otrzymywaniu rzadko występujących białek lub białek występujących w bardzo małych ilościach w materiałach biologicznych. Ilość, jakość i czystość otrzymanego białka rekombinowanego zależy od kilku czynników: rodzaju wektora ekspresyjnego i szczepu, doboru odpowiednich warunków wzrostu hodowli a także procedury oczyszczania.

Obecnie dostępnych jest szeroka gama systemów umożliwiających nadprodukcję rekombinowanych białek. Komórki eukariotyczne oraz prokariotyczne mogą być wykorzystywane jako fabryki do produkcji białek. Pierwszym krokiem jest odpowiedni dobór systemu ekspresyjnego. Należy uwzględnić warunki hodowli komórek, poziom ekspresji, umiejscowienie produkowanego białka oraz jego aktywność biologiczną.

Stosowanie zarazem komórki eukariotyczne jak i prokariotyczne ma swoje wady i zalety. Wykorzystanie komórek owadów lub ssaków umożliwia otrzymanie aktywnych cząsteczek białka ze wszystkimi potranslacyjnymi modyfikacjami. Wadą natomiast jest czasochłonna oraz kosztowna hodowla tych komórek, która prowadzi do otrzymania małej ilości preparatu. Dlatego komórki te zazwyczaj stosuje się wyłącznie do badań laboratoryjnych

Lepszym rozwiązaniem jest stosowanie komórek eukariotycznych. Komórki bakteryjne są częściej stosowane ze względu na znajomość mechanizmu transkrypcji i translacji oraz dostępność szerokiej gamy promotorów. Produkcja białek rekombinowanych w takich „fabrykach” jest również

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

szybką i tanią metodą. Bakterie jednak nie mają zdolności do przeprowadzania modyfikacji potranslacyjnych w produkowanych białkach co jest ich ogromną wadą.

Wybierając odpowiednie komórki do nadprodukcji białka warto spojrzeć również oczami krystalografa. Zastępując resztę metioniny w sekwencji białka rekombinowanego selenometioniną (Se-Met) wprowadzamy do struktury ciężki atom („znacznik” anomalny). Taka modyfikacja pomaga w sprawnym rozwiązaniu problemu fazowego. Selenometionina zostaje wprowadzona do komórek bakteryjnych wraz z pożywką, i w trakcie produkcji białka zastępuje reszty metioniny.

### **Oczyszczanie białka składa się z kilku głównych etapów:**

1. Izolacja docelowego białka z materiału biologicznego: rozbicie tkanek i komórek (sonifikacja, homogenizacja, zastosowanie wysokiego ciśnienia), ekstrakcja białka do roztworu wodnego.
2. Wstępne oczyszczanie (dodatkowe): wysolenie, dializa.
3. Oczyszczanie przy użyciu technik chromatograficznych (chromatografia powinowactwa, filtracja żelowa, chromatografia oddziaływań hydrofobowych).
4. Zmiana warunków jonowych (dodatkowe): dializa.
5. Pomiar stężenia i analiza czystości preparatu: elektroforeza.

Ważnym parametrem w trakcie oczyszczania białka jest temperatura, zazwyczaj proces ten jest prowadzony w niskiej temperaturze 2 - 4°C.

Chromatografia jest wciąż rozwijającą się metodą analityczną o coraz to większym zastosowaniu. Jest fizykochemiczną metodą rozdzielania składników jednorodnych mieszanin, wykorzystując różnice w ich podziale między fazą ruchomą i nieruchomą układu chromatograficznego. Składniki mieszaniny muszą różnić się właściwościami fizykochemicznymi, takimi jak: wielkością molekuł, ładunkiem, hydrofobowością, specyficznością substratową, co determinuje wybór odpowiedniego typu chromatografii. Metody chromatograficzne można wykorzystać do rozdziału różnych mieszanin związków organicznych (aminokwasów, białek, cukrów, lipidów, kwasów nukleinowych). Uwzględniając różne rodzaje fazy nieruchomej wyróżnić możemy chromatografię gazową, cieczową, bibułową i cienkowarstwową. W chromatografii cieczowej podczas rozdziału białek fazą ruchomą jest ciecz, a fazą nieruchomą jest ciało stałe, złożo hydrofilowe np. nierozpuszczalne związki agarozy, polimerów akryloamidu lub związki dekstranu. Chromatografię cieczową dzieli się na chromatografię kolumnową i planarną. W przypadku oczyszczania białek używa się tej pierwszej. W metodzie tej zazwyczaj używa się szklanej rurki, która wypełniona jest odpowiednim złożem. Złożo w kolumnie należy przemyć i wypełnić buforem o odpowiednim składzie i pH. Mieszaninę białek umieszczoną w buforze fazy ruchomej umieszcza się na wcześniej przygotowanej kolumnie.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Chromatograficzne techniki oczyszczania białek

### Filtracja żelowa

Dokonując rozdziału metodą filtracji żelowej (Gel Filtration, GE) uwzględnia się wielkość (masę) rozdzielanych białek. Metoda ta opiera się na zjawisku dyfuzji. Kolumna do chromatografii wypełniona jest ziarnami o porowatej strukturze. Analizując mieszaninę białek różniących się wielkością zauważymy wypływanie w pierwszej kolejności białek o największej masie. Białka te są za duże aby wejść w wąskie kanaliki porowatego wypełnienia, dlatego szybko przepływają przez kolumnę. Następnie kolumnę opuszczają białka średnich rozmiarów. Najwolniej przemieszczają się małe cząsteczki białek, wchodząc w wąskie kanaliki, dlatego wypływają jako ostatnie.

Kolumny z wypełnieniem do filtracji żelowej są dostępne komercyjnie jak i można je wypełniać samodzielnie. Najczęściej stosowane złoża, takie jak Sepharose, Sephadex lub Superdex, wykonane są z następujących materiałów; dekstran, bisakrylamid lub agaroz. Przygotowane własnoręcznie kolumny nie zawsze udaje się dobrze upakować, dlatego lepszym rozwiązaniem jest kupienie już gotowej. Zaletami omawianej metody jest łatwa obsługa i analiza otrzymanych wyników. Rozdział białek może zachodzić bez ich denaturacji. Warty podkreślenia jest fakt, że do nanoszenia preparatu jaki i wmywania białka można użyć tego samego buforu. Można rozdzielać mieszaninę nawet bardzo dużych cząsteczek białka i rozdział nie zależy od rodzaju użytego buforu. Omawiana metoda posiada również pewne ograniczenia. Cechuje się niską zdolnością rozdzielczą a objętości nanoszonych próbek są niewielkie. W przypadku dużej ilości preparatu, proces ten należy powtórzyć kilkakrotnie.

### Chromatografia jonowymienna

W chromatografii jonowymiennej (Ion Exchange Chromatography, IEC) rozdział następuje ze względu na wypadkowy ładunek białek. Kolumny mogą być wypełnione złożem obdarzonym dodatnim ładunkiem (anionit) lub ładunkiem ujemnym (kationit). Gdy kolumna wypełniona jest anionitem, silniej wiązane są białka naładowane ujemnym ładunkiem, które będą wypływać z kolumny jako ostatnie. Natomiast białka naładowane dodatnio lub obojętne, nie wiążą się ze złożem. Aby wmyć białka z kolumny stosowany jest gradient soli lub pH. W metodzie tej można również rozdzielać białka naładowane dodatnio, które różnią się wielkością ładunku. Analogicznie wygląda oczyszczanie preparatu białkowego w przypadku zastosowania kolumny wypełnionej kationitom. Kolumny są dostępne komercyjnie (HiTrap, HiLoad) lub można je samemu przygotować. Do wypełnienia kolumn stosowane są polimery węglowodanowe, takie jak agaroz, dekstran lub celuloza. W przypadku złoża naładowanego dodatnio, polimery te posiadają dodatkowo przyłączone łańcuchy węglowodorowe, które zakończone są dodatnimi grupami funkcyjnymi ( $-NHR_2$  lub  $-NR_3$ ). W przypadku kationity są to ujemnie naładowane grupy funkcyjne ( $-COOH$ ,  $-PO_4^{3-}$ ). Chromatografia jonowymienna cechuje się wieloma zaletami. Umożliwia rozdział próbek o dużej

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

objętości ale niskim stężeniu, ponieważ stosowane są kolumny o dużej pojemności. Jest to bardzo selektywna metoda i charakteryzuje się wysoką rozdzielczością. Do wad omawianej metody można zaliczyć konieczność nanoszenie próbki w buforze o określonym pH i dużą zawartość soli we frakcjach uzyskanych po rozdziale.

## Chromatografia oddziaływań hydrofobowych

W chromatografii oddziaływań hydrofobowych (Hydrophobic Interaction Chromatography, HIC) białka rozdzielane są ze względu na różnice w hydrofobowości. Kolumny stosowane w tej metodzie wypełnione są np. agarozą, polimerami akrylamidowymi, do których przyłączone są ligandy zakończone grupami hydrofobowymi (grupy fenyłowe). Rozdział dokonuje się dzięki adsorpcji białek na powierzchni złoża, poprzez hydrofobowe oddziaływania między białkiem a hydrofobowymi ligandami wypełnienia. Przed rozpoczęciem procesu rozdzielania, kolumnę należy przemyć roztworem o dużej zawartości soli, a następnie nanosi się badany preparat. Do wymywania stosowane są wodne roztwory soli o malejącym stężeniu bądź roztwory zawierające niewielką ilość związków organicznych. Wysoka jest również rozdzielczość tej metody.

Chromatografia oddziaływań hydrofobowych jest bardzo często stosowaną metodą. Tak, jak w przypadku chromatografii jonowymiennej, stosuje się kolumny o dużej pojemności, dzięki czemu można rozdzielać próbki o dużej objętości a małym stężeniu. Wadą tej metody jest obecność znaczącej ilości soli w końcowym preparacie, dlatego należy dodatkowo dializować próbkę lub poddać ją re-chromatografii stosując filtrację żelową.

## Chromatografia powinowactwa

Chromatografia powinowactwa (Affinity Chromatography, AC) jest najczęściej stosowaną metodą do izolowania i oczyszczania białek. Jest to chromatografia adsorpcyjna wykorzystująca wzajemne powinowactwo białka do liganda, który tworzy wypełnienie kolumny. Jako przykłady stosowanych układów białko – ligand, wymienić można: enzym – inhibitor, przeciwciało – antygen, receptor – hormon, białko wiążące maltozę – maltoza (MBP), znacznik histydynowy (His<sub>6</sub>) – jony niklu (II), S- transferaza glutationowa (GST) – glutation. Znacznik histydynowy, białko MBP i GST dołączone jest do oczyszczanego związku na etapie tworzenia białka rekombinowanego, tworząc w ten sposób białko fuzyjne. Po oczyszczeniu białka, wspomniane znaczniki należy usunąć. W tym celu stosuje się enzymy proteolityczne, prowadząc inkubację przez parę godzin a następnie jeszcze raz stosuje się chromatografię powinowactwa w celu usunięcia odciętych znaczników.

Po naniesieniu preparatu na kolumnę, w pierwszej kolejności wypływają z niej białka nie posiadające odpowiedniego znacznika, który umożliwił by ich związanie z kolumną. W kolejnym etapie kolumnę przemywa się roztworem zawierającym czynnik rozrywający wiązanie białko – ligand. Do wymywania można zastosować np. roztwór o dużym stężeniu maltozy w przypadku MBP, natomiast w przypadku kolumny wypełnionej jonami niklu (II) roztwór imidazolu. W ten

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

sposób możemy uzyskać czysty preparat, a do oczyszczenia którego użyliśmy tylko jednego etapu – jest to duża zaleta tej metody. Metodą tą można rozdzielać próbki o dużej objętości i różnym stężeniu. Umożliwia ona również zateżenie badanego materiału. Natomiast wadą tej metody jest trudność uzyskania niektórych specyficznych ligandów oraz konieczność przygotowywania kolumny we własnym zakresie.

Efekt każdego etapu procesu oczyszczania białka powinien być dokładnie monitorowany, co pozwoli nam znaleźć i wyeliminować ewentualne problemy mogące pojawić się w poszczególnych etapach procesu. Takie postępowanie umożliwi nam dopracowanie procedury oczyszczania docelowego preparatu. W tym celu powinna zostać określana ilość (stężenie) białka oraz jego aktywność. Po kolejnych etapach oczyszczania aktywność białka wzrasta, ponieważ powinno go być więcej w stosunku do innych białek i zanieczyszczeń.

### **Wykład 3: Krystalizacja białek**

#### **Białko wykorzystywane w procesie krystalizacji**

Jeśli planujemy wykorzystać białko do krystalizacji powinno być ono dostatecznie doczyszczane metodami chromatograficznymi. Roztwór białka musi zawierać odpowiedni bufor (np. Tris, HEPES), który zapewnia stałe pH oraz azydek sodu, który wykorzystywany jest do zabezpieczania roztworów przed rozwojem mikroorganizmów. Jeżeli roztwór białka powinien zawierać inny bufor niż ten po procesie oczyszczania, można go wymienić stosując np. dializę lub „odsalającą” kolumnę chromatograficzną. Białko powinno być homogenne i czyste w co najmniej 95% oraz odpowiednie stężenie. W razie potrzeby roztwór można zateżyć, stosując specjalne koncentratory z półprzepuszczalną błoną. Białko należy przechowywać w małych porcjach w temperaturze  $-70^{\circ}$ . Preparatów białkowych nie wolno wielokrotnie zamrażać i odmrażać.

#### **Pomiar stężenia białka**

Obecnie znanych jest wiele metod umożliwiających pomiar stężenia białka. Roztwór białka powinno zateżać się do wysokiego stężenia, należy jednak pilnować aby nie zaczęło się wytrącać. Roztwór białka zwykle ma stężenie 10-20mg/ml. Przed pomiarem stężenia należy preparat rozcieńczyć. Uwzględniając specyficzność próbki badanego roztworu można dobrać odpowiednią metodę spektrofotometryczną. Metody różnią się zakresem detekcji oraz czułością. Do oznaczania stężenia można np. wykorzystać: metodę pomiaru absorpcji w ultrafiolecie, metodą Bradforda lub metodę biuretową.

Obecnie najczęściej stosuje się metodę absorpcji UV. W metodzie tej wykorzystuje się absorpcję światła przy długości fali równej 280nm przez aminokwasy zawierające pierścień

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

aromatyczny (tryptofan, tyrozyna). Znając liczbę reszt tryptofanu ( $n_W$ ), reszt tyrozyny ( $n_Y$ ) a także jego masę molową białka ( $M_r$ ) można wyliczyć jego stężenie. Obecność kwasów nukleinowych może wpłynąć na błąd pomiaru, ponieważ absorbują światło przy  $\lambda=260\text{nm}$ . W przypadku roztworów białek i kwasów nukleinowych stosuje się metodę Warburga i Christiana.

Metoda Bradforda jest również często stosowana. Badane białko tworzy barwny kompleks z barwnikiem Coomassie Brilliant Blue (CBB), który wykazuje maksimum absorpcji przy długości fali równej  $595\text{nm}$ . Natężenie otrzymanej barwy kompleksu zależy od zawartości białka w roztworze.

Metoda biuretowa może być wykorzystana do pomiaru stężenia roztworów białek jak i różnych związków organicznych, które w swojej budowie posiadają wiązania peptydowe. Po dodaniu do badanych roztworów silnej zasady (NaOH) oraz siarczanu miedzi (II), następuje zmiana zabarwienia z niebieskiej na fioletową ( $\lambda=540\text{nm}$ ), ponieważ tworzy się związek kompleksowy jonów miedzi (II) z co najmniej dwoma wiązaniami peptydowymi.

## Właściwości kryształów białek

Kryształy białek mogą zawierać od 40-70% roztworu macierzystego (np. wody), co powoduje że mają odmienne właściwości w porównaniu do kryształów innych substancji. Pracując z kryształami białkowymi musimy być ostrożni i delikatni. Taka ilość wody powoduje, że cząsteczki białka są od siebie oddalone, ale zarazem są na tyle blisko aby mogły pojawić się oddziaływania, które umożliwią utworzenie uporządkowanej struktury krystalicznej. Otrzymane uporządkowanie nie jest aż tak doskonałe, co powoduje ograniczony zakres dyfrakcji. Obecność kanałów wodnych w kryształach białkowych może mieć zalety oraz wady.

Tak duża ilość roztworu macierzystego w kryształach gwarantuje, że cząsteczki białka przyjmują natywną konformację, podobną do tej obserwowanej w warunkach fizjologicznych. Obecność kanałów wodnych również w łatwy sposób umożliwia wprowadzenie do kryształów białkowych składników małocząsteczkowych, np. jonów ciężkiego pierwiastka, inhibitorów lub aktywatorów.

Obecność wody w kryształach jest również negatywnym zjawiskiem. W trakcie pomiaru rentgenograficznego, podczas naświetlania jonizującym promieniowaniem wytwarzane są wolne rodniki. Dzięki kanałom wodnym te bardzo niebezpieczne drobiny w szybki sposób rozprzestrzeniają się, docierając do cząsteczek białka i powodując ich niszczenie, degradację. Można jednak ten negatywny proces w dużej mierze zatrzymać. Stosujemy w tym celu np. zjawisko „witryfikacji” roztworu rozpuszczalnika stosując niskie temperatury lub można wykorzystać wymiatacz wolnych rodników.

Praca z kryształami białek nie należy do łatwej. Kryształy te zachowują ograniczoną trwałość jedynie w wąskim zakresie temperatury, pH oraz siły jonowej. Jeśli zostaną pozostawione na powietrzu tracą wodę i od razu ulegają nieodwracalnym przemianom. Dlatego po wyjęciu z

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

roztworu macierzystego należy je od razu umieścić np. w ciekłym azocie. Są one bardzo miękkie, delikatne, praca z nimi wymaga dużego doświadczenia.

## Proces tworzenia kryształów białek. Wykres fazowy krystalizacji

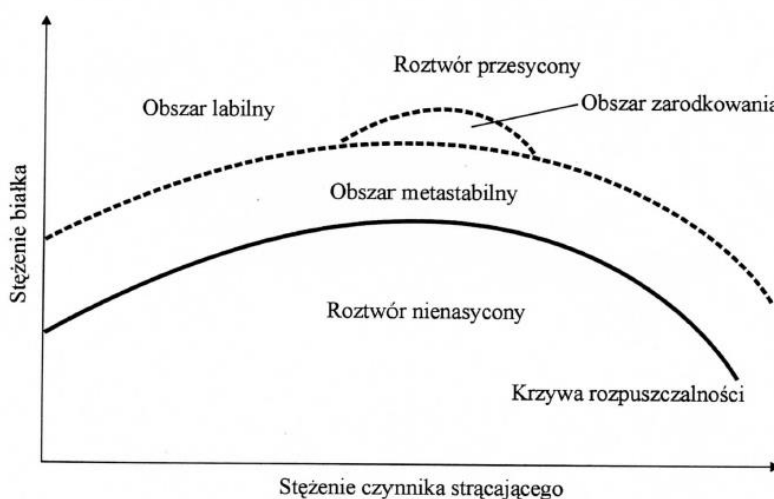
Warto w tym punkcie przypomnieć drugą zasadę termodynamiki, która mówi, że w dowolnym procesie w układzie termodynamicznie izolowanym entropia, czyli miara stopnia nieuporządkowania układu, nigdy nie maleje. Na tej podstawie można by było stwierdzić, że proces krystalizacji białek zachodzi wbrew II zasadzie dynamiki. Następuję przecież przejście od chaosu, nieuporządkowania do porządku. O tym czy proces zajdzie decyduje jeszcze inna funkcja stanu, a mianowicie entalpia swobodna Gibbsa. Na jej wartość wpływa entalpia, entropia oraz temperatura ( $G = H - TS$ ). Oddziaływania międzycząsteczkowe tworzące się w trakcie formowania kryształów mogą przeważać człon entropowy, co umożliwi zajście tego procesu. Ważnym rolę odgrywa też temperatura, która jest obecna we wzorze. Procek, który w wyższej temperaturze był niekorzystny, w niższej temperaturze może zachodzić spontanicznie. W związku z tym proces poleca się prowadzić proces krystalizacji białek w temperaturze 4°C.

Krystalizacja białek jest procesem egzotermicznym polegającym na powstaniu fazy krystalicznej z fazy ciekłej (roztworu wodnego). Proces krystalizacji zwykle składa się z czterech faz. Pierwsza faza to powstawanie zarodków krystalizacyjnych, czyli nukleacja. W kolejnym etapie następuje swobodny wzrost pojedynczych kryształów (monokryształów), a następnie reorganizacja warstwy powierzchniowej i powstawanie tzw. mikrostruktury krystalicznej. Końcowy etap to łączenie się pojedynczych kryształów w większe struktury. Stosując odpowiednie warunki w trakcie procesu możemy zatrzymać się na drugim etapie i otrzymać odpowiedni do badań krystalograficznych monokryształy.

Diagram fazowy dla procesu krystalizacji przedstawia Rysunek 3.1. Roztwór przesycony otrzymuje się w przypadku wysokich stężeń białka i czynnika strącającego, skutkuje to powstaniem amorficznych agregatów białkowych. Obniżenie stężenia może prowadzić do tworzenia się zarodków krystalizacyjnych, które mogą przemienić się w monokryształy. W obszarze metastabilnym również tworzą się zarodki ale ich wzrost nie jest spontaniczny. Umieszczenie kryształu białka w tym obszarze może powodować jego wzrost. Przy niższych stężeniach tworzy się roztwór nienasycony – białka są całkowicie rozpuszczalne.



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki



**Rysunek 3.1** Wykres fazowy procesu krystalizacji białka.

### Skład roztworu krystalizacyjnego

Podstawowymi składnikami roztworu krystalizacyjnego jest białko, woda, bufor i czynnik strącający. Poza tymi głównymi składnikami w roztworze może występować np. azydek sodu (czynnik chroniący przed rozwojem pleśni i bakterii), kofaktory (jony metali lub inne związki organiczne, dzięki którym enzym zdolny jest do pełnienia swojej funkcji). Ochronę białka przed utleniającym wpływem powietrza można zapewnić dodając do roztworu stosuje się różne środki redukujące. Można stosować ditiotreitol (DTT) lub merkaptoetanol, które zapobiegają utlenieniu grup tiolowych w białku, co chroni cząsteczkę przed zmianą konformacji oraz utratą jej aktywności. W celu usunięcia jonów metali, które mogą przeszkadzać w procesie krystalizacji wykorzystać można czynnik kompleksujący (np. EDTA)

### Czynniki strącające

Czynniki strącające można podzielić na trzy główne grupy: sole, rozpuszczalniki organiczne i polimery. Siarczan (VI) amonu jest najczęściej stosowaną solą. Stosować również można inne sole nieorganiczne (fosforany, azotany, chlorki) jak i organiczne sole (mrówczany, octany, cytryniany). Do drugiej grupy czynników strącających zaliczyć przede wszystkim różnego typu alkohole: etanol, izopropanol czy MPD (2-metylo-2,4-pentanodiol). Trzecia grupa to polimery o różnej długości łańcuch, czyli różnej masie cząsteczkowej. Najczęściej stosowany jest glikol polietylenowy (PEG). Długość łańcucha zmienia właściwości polimerów, np. PEG-400 jest „lekki” a jego skrajne grupy wykazują charakter alkoholowy. Natomiast, PEG-20.000, jest „ciężki” a końcowe grupy OH nie wpływają na odczyn.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Czynniki wpływające na wzrost kryształów

Na wzrost kryształów mają wpływ różne czynniki. Do najważniejszych można zaliczyć stężenie białka, temperaturę, rodzaj i pH buforu, rodzaj i stężenie czynnika strącającego. Istnieje ogromna liczba innych czynników, chemicznych (jony metali, detergenty, siła jonowa), fizycznych (ciśnienie, dźwięk, wibracja) jak i biochemicznych (czystość i jednorodność próbki, modyfikacje sekwencji, obecność ligandów), które wpływają na ten proces.

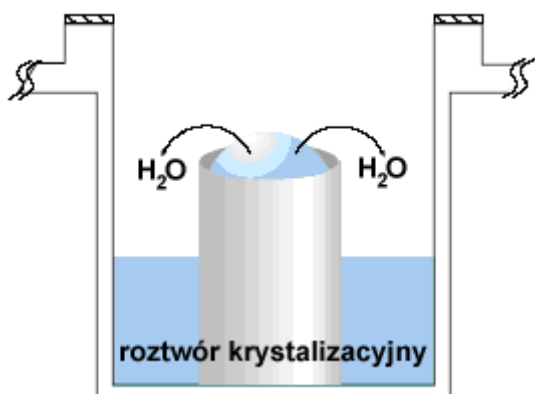
Tak ogromna ilość różnych czynników wpływających na wzrost kryształów sprawia, że dostaje się ogromną ilość warunków do przetestowania. Sprawdzenie wszystkich możliwości jest bardzo czasochłonne i w ręcz niemożliwe, ponieważ nie dysponuje się taką ilością preparatu. Jak sobie z tym problemem poradzić? Jeśli nie jest to unikatowe białko i należy do rodziny znanych białek, to można w literaturze poszukać warunków krystalizacji tych białek. Warunki te będą dobrym punktem startu w krystalizacji badanego białka. Jeśli w literaturze nie znajdziemy informacji na ten temat możemy skorzystać z komercyjnie dostępnych specjalnych zestawów odczynników tzw. „screenów”.

## Metody otrzymywania kryształów

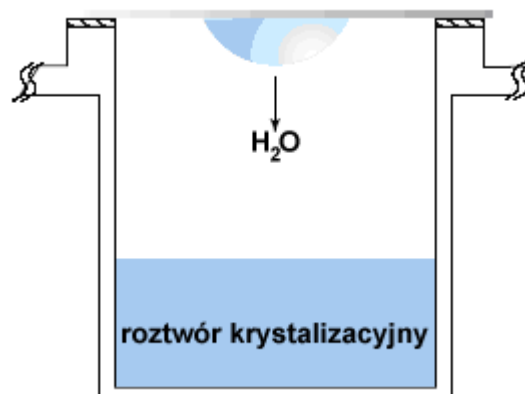
Kryształy białek zawierają 40-70% rozpuszczalnika (wody), dlatego struktura białek w roztworze (w komórce) jest bardzo podobna do struktury w kryształach. Dzięki temu, białka przyjmować mogą taką samą konformację jak w komórce. Dlatego, enzymy mogą nadal pełnić swoją funkcję oraz są zdolne do przeprowadzenia reakcji katalitycznej nawet w formie krystalicznej. Obecność tak dużej ilości wody utrudnia pracę z kryształami białkowymi, ponieważ stają się one bardzo delikatne, podatne na uszkodzenia i szybko ulegają rozpadowi po wyjęciu z roztworu macierzystego. Wiązania wodorowe, oddziaływania hydrofobowe i mostki solne między cząsteczkami białek odpowiedzialne są za formowanie i stabilizację kryształów białek.

Obecnie istnieje wiele metod otrzymywania kryształów, które można podzielić na dwie grupy. Pierwsza grupa to metody krystalizacji związane z dyfuzją przez parowanie, do których zaliczymy "hodowanie" w warunkach siedzącej kropli (Rysunek 3.2), wiszącej kropli (Rysunek 3.3) czy rzadziej spotykane uzyskiwanie kryształów w kropli kanapkowej. W kolejnej grupie znajdują się pozostałe znane metody, do których zaliczamy: dializę, mikrodializę, techniki batch i mikrobatches.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki



**Rysunek 3.2** Metoda siedzącej kropli.



**Rysunek 3.3** Metoda wiszącej kropli.

Najczęściej stosowanymi metodami są techniki oparte na dyfuzji przez parowanie, ponieważ umożliwiają uzyskanie kryształów w prosty sposób z niewielkiej ilości próbki badanego preparatu. W metodach tych przygotowuje się dwa roztwory, jeden jest bardziej a drugi mniej stężony (rozcieńczony przez dodatek białka). Roztwór krystalizacyjny (bardziej stężony) mogący zawierać sól, detergent, środek strącający i bufor, umieszcza się na dnie naczynia, Drugi roztwór zawiera oprócz roztworu krystalizacyjnego również białko, tworzy on krople, w której będą tworzyć się kryształy. Ponieważ roztwory różnią się stężeniem zachodzi powolne odparowanie roztworu z kropli, następuje zatężenie tego roztworu to może sprzyjać formowaniu się kryształów.

W metodzie „siedzącej kropli” (Rysunek 3.2) kropla białko-roztwór krystalizacyjny umieszczona jest w specjalnym dołku, natomiast roztwór bardziej stężony (krystalizacyjny) znajduje się na dnie naczynia. Oba roztwory nie mają bezpośredniego kontaktu ze sobą. Wadą tej techniki może być adhezja kryształów białka do powierzchni tworzywa, z którego wykonane są płytki do krystalizacji. Proces ten może utrudnić wyjęcie kryształu z kropli. Czasami jednak, zjawisko adhezji, może inicjować formowanie kryształów

W technice „wiszącej kropli” (Rysunek 3.3) krople białko-roztwór macierzysty nanosi się na okrągłe szkiełko pokryte zazwyczaj silikonem, co zapobiega adhezji kryształów białka. Szkiełko z kroplą białka odwraca się do góry dnem. W ten sposób kropla wisi nad roztworem krystalizacyjnym umieszczonym na dnie naczynia. Szklane szkiełko pozwala na monitorowanie postępów formowania się kryształów,

## Metody szczepienia krystalizacji

Można wyróżnić dwie metody szczepienia: makroszczepienie i mikroszczepienie. Techniki te można stosować, jeśli otrzymane wcześniej kryształy są złej jakości. Mikroszczepienie polega na umieszczeniu zarodków krystalizacyjnych w kropli w stanie metastabilnym. W tym celu używa się

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

włosa (np. kociego) lub końcówki pipety, którą zanurza się w rozworze z zarodkami a następnie niewielka ilość zarodków przenosi się do nowej kropli. W nowej kropli po jakimś czasie z zarodków zaczynają tworzyć się kryształy. Gdy dysponujemy większymi kryształami ale są one źle uformowane, możemy je również wykorzystać do mikroszczepienia. Kryształy zostają rozdrobnione i tworzy się z nich zawiesinę. Następnie przygotowujemy serię rozcieńczonych roztworów. Tak przygotowanymi zawiesinami szczepi się zrównoważone krople.

Druga metoda polega na przeniesieniu przemytego pojedynczego monokryształu z kropli, w której już nie rośnie do nowej kropli.

## Roboty do krystalizacji

Roboty do krystalizacji są bardzo użytecznym narzędziem w procesie krystalizacji białek. Umożliwiają w szybszy i dokładniejszy sposób sprawdzić różne warianty warunków krystalizacji badanego białka. Ważną zaletą jest możliwość używania bardzo małych objętości próbki (nawet 125  $\mu$ l), co umożliwia oszczędzenie białka i pozwala przetestować większą liczbę warunków. Wyniki krystalizacji są fotografowane co 8 godzin, co usprawnia proces monitorowania procesu.

## Wykład 4: Pomiar synchrotronowy

Metody dyfraktometryczne obecnie znajdują szereg zastosowań w różnych dziedzinach nauki. Wykorzystywane są w fizyce ciała stałego, badaniach materiałowych oraz krystalografii. Dyfraktometria rentgenowska jest przede wszystkim wykorzystywana do badań naukowych w celu określenia dokładnej struktur 3D różnych substancji, związków małowcząsteczkowych jak i biomakromolekuł.

Metoda ta bazuje na oddziaływaniu promieniowania X z chmurami elektronowymi poszczególnych atomów wchodzących w skład badanego związku. W skutek tego oddziaływania otrzymuje się obrazy dyfrakcyjne. Z pomiaru dyfraktometrycznego otrzymuje się informacje o parametrach komórki elementarnej oraz intensywności wiązki odbitej (refleksów). Analizując otrzymane obrazy, uzyskane wskutek dyfrakcji promieni X na monokryształach, wyznaczona zostaje trójwymiarowa mapa gęstości elektronowej. Analiza otrzymanej mapy dostarcza szereg przydatnych informacji, pozwala:

- wyznaczyć położenie atomów w sieci krystalicznej,
- wyznaczyć geometrię walencyjną (długości wiązań i wartości kątów) oraz konformację (kąty torsyjne),
- wyznaczyć wewnątrz- i międzycząsteczkowe oddziaływania,
- obliczyć moment dipolowy wiązań, cząsteczek,
- ustalić konfigurację absolutną,
- wyznaczyć położenie i odległości pomiędzy cząsteczkami w sieci krystalicznej,

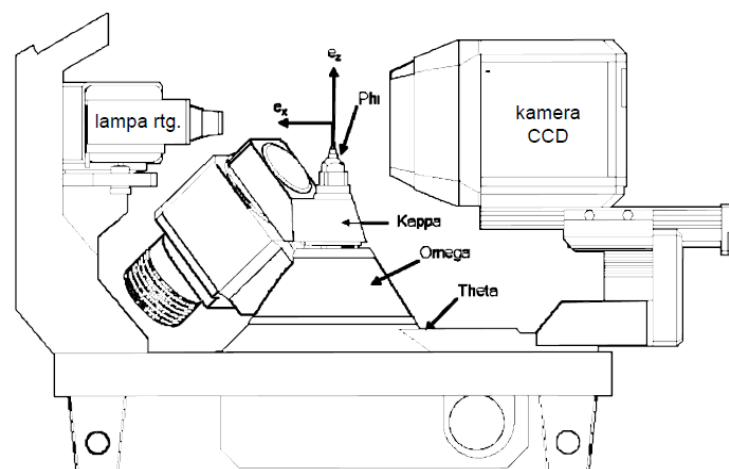
Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Typowy eksperyment dyfrakcyjny wymaga trzech rzeczy: źródła promieniowania rentgenowskiego, kryształu oraz detektora wiązek ugiętych. Na nowoczesnych liniach pomiarowych znajdziemy znacznie więcej aparatury, np. elementy optyki (monochromatory, zwierciadła, kolimatory, etc.), komputery, roboty, przystawki niskotemperaturowe, kamery video, mikroskopy czy układy zabezpieczeń, ale te trzy podstawowe elementy będą zawsze obecne.

Obiektem w badaniach strukturalnych jest monokryształ badanego związku, który musi spełniać kilka ważnych kryteriów. Kryształ powinien być otoczony naturalnymi ściankami, posiadać wymiary mniejsze od średnicy wiązki promieniowania oraz być trójwymiarowy (rozmiar w każdym zakresie 0,2 – 0,5mm). Kryształy o większych rozmiarach można spróbować przeciąć np. skalpelem, ale należy robić to bardzo delikatnie, żeby nie naruszyć struktury kryształu. Badany kryształ powinien być trwały w trakcie trwania całego pomiaru oraz nie może posiadać żadnych pęknięć. Kryształy, które są nie trwałe w temperaturze pokojowej (np. kryształy białkowe) powinny być mierzone w niskiej temperaturze (np. pod strumieniem ciekłego azotu). Kryształy nieodporne na wilgoć można obtoczyć w smarze, oleju parafinowym albo montować w specjalnych kapilarach. Wyboru odpowiedniego monokryształu dokonuje się przy użyciu polaryzacyjnych mikroskopów. Pomiar monokryształu, spełniającego powyższe warunki, wykonuje się za pomocą dyfraktometru rentgenowskiego (Rysunek 4.1).

#### Budowa typowego dyfraktometru rentgenowskiego:

1. źródło promieni rentgenowskich: generator wysokiego napięcia + lampa rentgenowska,
2. goniometru pozwalający dokładne ustawienie kryształu,
3. detektora,
4. komputera sterującego przebiegiem pomiaru i obróbką danych.



**Rysunek 4.1** Schemat dyfraktometru rentgenowskiego.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Kolejnym etapem po wyborze odpowiedniego do badań monokryształu jest jego zamocowanie na goniometrze, który pozwala zmieniać jego orientację w trakcie montowania oraz przebiegu pomiaru. Kryształ na główce goniometrycznej można zamontować różnymi sposobami. Związki małowcząsteczkowe można przykleić na pręcik szklany główki goniometrycznej za pomocą żywicy chemoutwardzalnej lub lakieru. Nietrwale i wrażliwe substancje można umieścić w szklanej kapilarze lub mocować przy użyciu oleju silikonowego lub parafinowego.

Badanie struktury białka metodą dyfrakcji promieni rentgenowskich wymaga pojedynczego kryształu uformowanego z badanego materiału. Krystalizacja białek jest jednym z trudniejszych etapów rozpoczynających proces rozwiązywania struktury biomakromolekuły. Kryształ białka przeznaczony do pomiarów dyfrakcyjnych jest z reguły jeszcze „dodatkowo zamrożony” w parach ciekłego azotu o temperaturze  $-170^{\circ}\text{C}$ . Ma to na celu zabezpieczenie wrażliwego białka i towarzyszącej mu zawsze w kryształach otoczki wodnej.

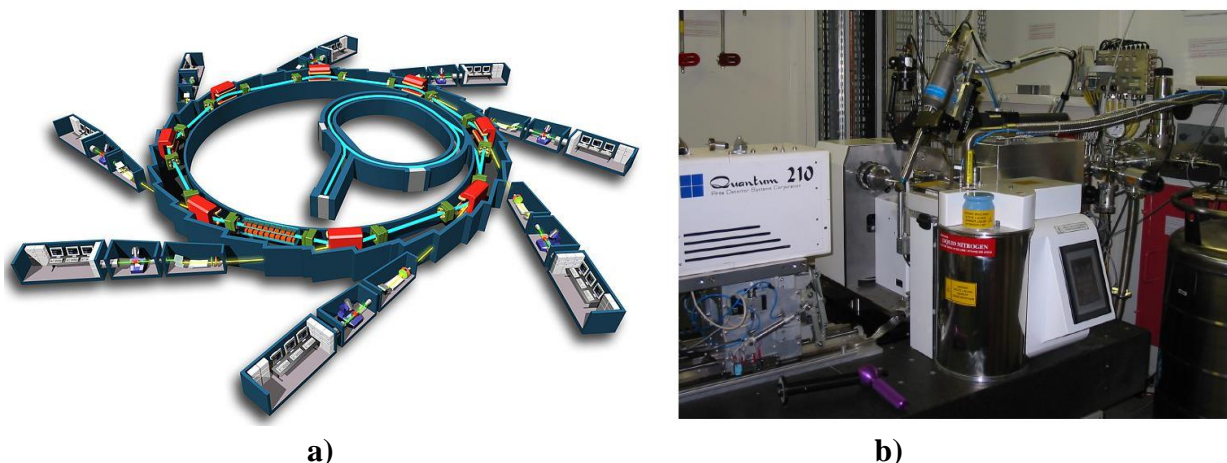
Ze względu na szybką destrukcję na skutek dehydratacji, kryształy białek wyjęte z roztworu krystalizacyjnego (macierzystego) muszą być natychmiast zabezpieczone przed utratą wody. Do niedawna główny sposób polegał na zamknięciu ich w cienkościennej kapilarze szklanej w obecności kropelki roztworu macierzystego. Metody badań dyfrakcyjnych w warunkach kriogenicznych (w bardzo niskiej temperaturze) podpowiedziały lepszy sposób. Kryształy "montuje" się obecnie w małych pętelkach z włókna nylonowego, którymi wyławia się je z roztworu, po czym pętelkę z kryształem błyskawicznie umieszcza się w strumieniu gazowego azotu o temperaturze 100 K ( $-173^{\circ}\text{C}$ ). Zabieg ten powoduje witrifikację (zeszklenie) obecnej w kryształach wody. Dzięki temu unika się dyfrakcji od kryształów lodu oraz utrudnia wolnym rodnikom penetrację kryształu. Korzystne w tej metodzie jest i to, że ilość roztworu pobrana wraz kryształem jest minimalna: tyle tylko ile potrzeba do utworzenia błonki powierzchniowej wewnątrz pętelki (o rozmiarach ułamka milimetra), w której zawieszono kryształ. Problemem w tej metodzie jest czasami obecność soli w roztworze macierzystym, które podczas chłodzenia mogą gwałtownie wykryształizować uniemożliwiając pomiary dyfrakcji od kryształu białka. Zapobiega się temu dodając do roztworu macierzystego związki "przeciwzamrożeniowe" zwane krioprotektantami, podobne do płynów dodawanych zimą do chłodnic samochodowych. Popularnymi krioprotektantami są np. glicerol i glikol etylenowy. Ciekawą demonstracją szkód dokonanych w kryształach białka przez wolne rodniki jest obserwacja eksplozji kryształu naświetlonego promieniowaniem rentgenowskim w momencie wyłączenia urządzenia chłodzącego. Eksplozja wywołana jest gwałtownym uwolnieniem gazów powstałych w reakcjach rodnikowych, głównie  $\text{CO}_2$  (dekarboksylacja grup  $\text{COO}^-$ ) i  $\text{H}_2$  (rodnikowy rozkład wody).

Tradycyjnym źródłem promieni Roentgena jest lampa rentgenowska, tj. rura próżniowa, w której rozpędzone do energii kilkudziesięciu tysięcy eV elektrony bombardują anodę-tarczę wykonaną z miedzi lub z molibdenu. Energia wyhamowanych elektronów przede wszystkim rozgrzewa anodę, a tylko część emitowana jest w postaci kwantów promieniowania

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

elektromagnetycznego. Niewielka część energii pochłaniana jest przez atomy tarczy, jonizując ich najgłębszą powłokę elektronową K. Zapelnienie tej wakancji przez elektrony kolejnej powłoki wiąże się z emisją fotonu oznaczanego jako  $K\alpha$ . W przypadku miedzi, to charakterystyczne promieniowanie  $Cu K\alpha$  ma długość  $1.5418 \text{ \AA}$ , bardzo użyteczną do badań dyfrakcyjnych [ $\lambda(\text{Mo } K\alpha) = 0.7107 \text{ \AA}$ ]. Problem niskiej intensywności takich lamp częściowo rozwiązano konstruując generatory z maszyną wirującą anodą.

W badaniach biokrytalograficznych korzysta się jednak obecnie głównie z promieniowania rentgenowskiego generowanego w synchrotronach (Rysunek 4.2). Synchrotron jest cyklotronem o średnicy setek metrów, w którym po torze zamkniętym krążą elektrony z relatywistyczną prędkością. Synchrotrony składają się z pierścieniowego tunelu wypełnionego próżnią. Na obwodzie tunelu rozmieszczone są potężne elektromagnesy tworzące pole magnetyczne, które wpływa na zakrzywienie toru ruchu elektronów. Elektrony są wstrzyknięte z akceleratora liniowego do tunelu, następnie zostają przyspieszane. Krążące w próżni elektrony mają stałą energię, a następnie ulegają przyspieszeniu kątowemu na skutek przelatywania przez pole magnetyczne. Źródłem pola magnetycznego o szerokim zakresie długości fali, od mikrofal do twardego promieniowania rentgenowskiego są elektromagnesy zakrzywiające.



**Rysunek 4.2** a) Schemat ośrodka synchrotronowego, b) Stacja pomiarowa ID14 do badań krytalograficznych makromolekuł synchrotronu ESRF w Grenoble.

Pomiędzy elektromagnesami zakrzywiającymi w odcinkach prostoliniowych mogą zostać umieszczone dodatkowe urządzenia (ondulatory), które mogą wywołać dodatkowe przyspieszenie kątowe przelatujących elektronów. Zwiększając w ten sposób intensywność promieniowania synchrotronowego. Otrzymujemy w ten sposób promieniowanie silniejsze niż gdy wykorzystamy tylko magnesy zakrzywiające.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

### Do głównych zalet promieniowania synchrotronowego można zaliczyć:

- wysoką intensywność (wiele rzędów wyższa niż w lampach rentgenowskich),
- szeroki zakres widmowy,
- możliwość dostrojenia długości fali do potrzeb pomiaru,
- polaryzację,
- strukturę czasową (promieniowanie emitowane jest w "pęczkach" odpowiadających pakietom krążących w pierścieniu synchrotronowym elektronów).

Największymi na świecie źródłami promieniowania synchrotronowego jest ośrodek APS (Advanced Photon Source) znajdujący się pod Chicago w USA i ośrodek Spring-8 leżący niedaleko Osaki w Japonii. Natomiast w Europie największe synchrotrony znajdują się w Grenoble (ESRF, European Synchrotron Radiation Facility) i w Hamburgu (DESY, Deutsches Elektronen-Synchrotron). Również Polska od 2015 roku dysponuje synchrotronem, który został wybudowany w Krakowie przez Narodowe Centrum Promieniowania Synchrotronowego Solaris.

### **Wady i zalety dyfraktometrii rentgenowskiej**

W trakcie pomiaru rentgenowskiego próbka badanego związku nie ulega zniszczeniu, jest to główną zaletą tej metody. Możliwe jest zatem ponowne wykorzystanie próbki do innych badań. Wiele innych używanych metod ilościowych i jakościowych powoduje zniszczenie materiału badanego. Jest to bardzo ważne podczas badania wyjątkowych, unikatowych materiałów np. medycznych lub archeologicznych.

Badania strukturalne umożliwiają również bardzo dokładne ustalenie struktury przestrzennej badanych związków małowcząsteczkowych jak i biomakromolekuł. Możliwe jest także ustalenie konfiguracji absolutnej związków optycznie czynnych, co jest istotną informacją dla chemika organika. Tak naprawdę, żadna inna metoda analityczna nie daje takiej pełnej informacji odnośnie struktury 3D badanych związków.

Rentgenografia strukturalna ma również wady. Materiałem do badań musi być bardzo dobrej jakości monokryształ, które nie zawsze jest łatwy do uzyskania na drodze krystalizacji. Badany kryształ powinien mieć również odpowiednie wymiary, być trójwymiarowy i nie może posiadać żadnych pęknięć. Monokryształ powinien być także trwały w temperaturze pokojowej i odporny na działanie promieniowania rentgenowskiego. Takie ograniczenie można rozwiązać wykorzystując dyfraktometrię z przystawkami niskotemperaturowymi.

Wadą rentgenografii jest także stosunkowo wysoki koszt i czasochłonność wykonywania takiego pomiaru. Koszt zakupy dyfraktometru rentgenowskiego wynosi w granicach 100–500 tys. euro. Czas pomiaru może wynosić od kilku godzin dla związku metaloorganicznego do nawet kilku tygodni w przypadku biomakromolekuł. Rozwiązanie struktury w miarę prostego związku jest dość szybkie, podczas gdy określenie struktury dużych cząsteczek np. białek może zająć do kilku tygodni.



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Bardzo poważną wadą rentgenografii jest stosowanie silnego promieniowania rentgenowskiego. Osoby pracujące przy dyfraktometrze są na nie narażone, co może bardzo negatywnie wpływać na ich zdrowie a nawet może grozić utratą życia. Obecnie używane aparaty sterowane są przy pomocy komputera oraz wyposażone są w specjalne osłony, które zatrzymują promieniowanie rentgenowskie. Posiadają również specjalne zabezpieczenia, dzięki którym nie można rozpocząć pomiaru bez zamknięcia obudowy.

## Wykład 5: Rozwiązywanie struktur krystalicznych

W trakcie pomiaru dyfraktometrycznego promieniowanie rentgenowskie rozpraszane jest przez elektrony. Liczba atomowa atomu rozpraszającego oraz kąt rozproszenia ( $2\theta$ ), wyrażonego atomowym czynnikiem rozpraszania  $f$ , wpływają na stopień tego rozproszenia. W przypadku grupy atomów amplitudę i fazę promieniowania rentgenowskiego rozpraszanego przez komórkę podstawową wyraża się za pomocą czynnika struktury:

$$F(hkl) = A(hkl) + iB(hkl)$$

dla każdego odbicia.

Można go obliczyć za pomocą równań:

$$A(hkl) = \sum_j f_j \cdot \cos 2\pi(hx_j + ky_j + lz_j)$$

$$B(hkl) = \sum_j f_j \cdot \sin 2\pi(hx_j + ky_j + lz_j)$$

przy czym sumowania dokonuje się po wszystkich atomach w komórce elementarnej. Fale można przedstawić za pomocą liczb zespolonych:

$$A_j + iB_j = a_j \cdot e^{i\alpha_j}$$

$e_j$  - amplituda,  $\alpha_j$  - kąt fazowy.

Odwzorowanie to stanowi dogodny sposób przedstawienia dwóch prostopadłych składowych wektora (dla  $0^\circ$  i  $90^\circ$ ) w jednym równaniu (A jest składową dla  $0^\circ$ , a B dla  $90^\circ$ ).

Celem pomiaru jest uzyskanie trójwymiarowej mapy gęstości elektronowej, której interpretacja pozwala ustalić położenie atomów. Aby to osiągnąć należy wykonać trójwymiarowe sumowanie Fouriera. Szeregi Fouriera stosuje się ze względu na możliwość wykorzystania ich do opisu periodycznych funkcji i periodycznego rozmieszczenia rozpraszającej materii tworzącej kryształ. Liczbę elektronów przypadającą na jednostkę objętości lub gęstość elektronową w danym punkcie  $x,y,z$  oznaczoną symbolem  $\rho(x,y,z)$ , podaje wyrażenie:

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

$$\rho(x, y, z) = 1/V_k \sum \sum \sum F(hkl) \exp\{-2\pi i(hx + ky + lz)\}$$

$V_k$  - objętość komórki elementarnej,

$F(hkl)$  - czynnik struktury dla wskaźników  $hkl$ .

Uwzględniając prawidłowość struktury i zakładając, że:

$$\sin(-x) = -\sin(x) \quad ; \quad \cos(-x) = \cos(x);$$

$$F_{hkl} = F_{-h-k-l}$$

część wyrażen na gęstość elektronową znosi się i otrzymujemy:

$$\rho(x, y, z) = (|F(000)|)/V_k + Z/V_k \sum \sum \sum |F| \cos(\phi - \alpha)$$

Gdybyśmy znali wartość  $F$  i  $\phi$  dla każdego  $hkl$  moglibyśmy obliczyć  $\rho$  dla wszystkich wartości  $x, y, z$  uzyskując trójwymiarową mapę gęstości elektronowej. Jednak dane eksperymentalne dostarczają tylko modułu  $|F|$  obliczonego na podstawie natężenia  $I$ , nie dostarczają natomiast fazy tej funkcji zespolonej. Dlatego też problem fazowy jest podstawowym problemem, jaki należy rozwiązać w analizie strukturalnej. Istotę problemu fazowego obrazuje diagram Hosemanna i Bagchiego (Rysunek 5.1).

$$\begin{array}{ccc}
 I(b) \leftrightarrow FF^*(b) & \xleftrightarrow{\quad} & F(b) \\
 \updownarrow & & \updownarrow \\
 P(u) & \xleftrightarrow{\quad} & \rho(r)
 \end{array}$$

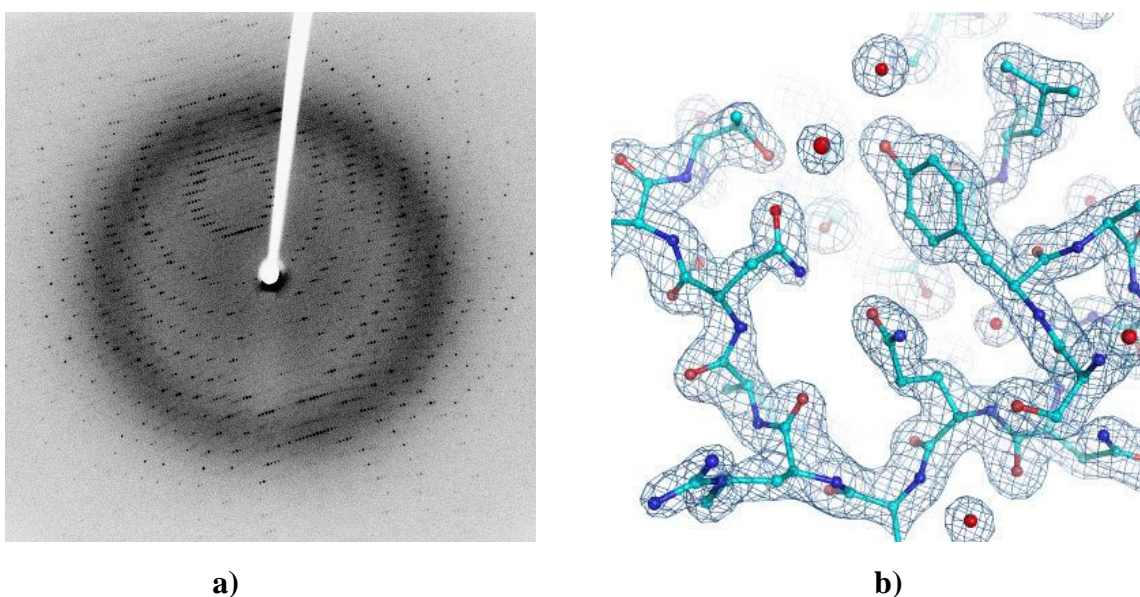
**Rysunek 5.1** Diagram Hosemanna i Bagchiego, w którym  $I(b)$  oznacza natężenie wiązki ugiętej,  $F(b)$  – czynnik struktury,  $P(u)$  – funkcja Pattersona i  $\rho(r)$  – gęstość elektronowa.

Etapy możliwe do wykonania oznaczone są linią ciągłą, natomiast linią przerywaną zaznaczone są etapy, w których wymagana jest znajomość fazy czynników struktury. Rozwiązanie problemu fazowego dla związków małowartościowych można znaleźć poprzez interpretację funkcji Pattersona  $P(u)$  lub metodami bezpośrednimi. Podstawą metod bezpośrednich są oczywiste fakty, że gęstość elektronowa w żadnym miejscu nie może być mniejsza niż zero, oraz że ma ona rozkład "atomowy", tj. skoncentrowana jest głównie w rdzeniach atomowych struktury. Te proste warunki wprowadzają tak silne ograniczenia na zależności między czynnikami struktury, że udaje się ostatecznie wyprowadzić relacje wiążące ich fazy. Wybór jednej z metod uzależniony jest w dużej mierze od składu chemicznego analizowanego kryształu. Funkcję Pattersona można stosować w przypadku struktur zawierających tzw. atomy ciężkie o dużej liczbie atomowej.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Niestety o ponad dwa rzędy wielkości większe struktury kryształów makromolekuł nie dają się rozwiązać metodami bezpośrednimi. Należało opracować inne. Pionierem na tym polu był Max Perutz, który pracując nad strukturą hemoglobiny stworzył w latach 50. metodę podstawienia izomorficznego (MIR). W latach 60., głównie dzięki pracom Michaela Rossmanna, powstała metoda podstawienia cząsteczkowego (MR). W latach 80. przyszła kolej na metodę dostrojonej dyfrakcji anomalnej (MAD), kojarzoną zwykle z nazwiskiem Wayne'a Hendricksona. Wszystkie trzy metody bazują w pewnym stopniu na funkcji Pattersona.

Rentgenowska analiza strukturalna monokryształów opiera się na zjawisku dyfrakcji promieniowania rentgenowskiego na kryształach. Na podstawie obrazu dyfrakcyjnego (Rysunek 5.2a) otrzymanego przez rozpraszanie promieni rentgenowskich na molekułach lub atomach periodycznie ułożonych w kryształach można obliczyć mapę rozkładu gęstości elektronowej (Rysunek 5.2b). Model struktury otrzymany przez interpretację rozkładu gęstości elektronowej jest rozbudowywany i udokładniany z użyciem danych z pomiaru dyfrakcyjnego, co pozwala na otrzymanie dokładnej struktury badanego kryształu.



**Rysunek 5.2** a) Przykładowy obraz dyfrakcyjny. b) Przykładowa mapa gęstości elektronowej.

Proces analizy rentgenowskiej monokryształów można podzielić na dwa etapy:

- wyznaczenie parametrów komórki elementarnej, klasy Lauego, typu sieci przestrzennej i układu krystalograficznego,
- wyznaczenie położenia atomów lub jonów w komórce elementarnej na podstawie pomiaru i analizy natężeń promieni ugiętych (refleksów).

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Wstępny pomiar rentgenograficzny (pre-experiment) umożliwia ocenę jakości wybranego do pomiaru monokryształu. Jeżeli jego „zdolność dyfrakcji” jest dobra oraz jest trwały podczas naświetlania promieniowaniem rentgenowskim, możemy przejść do pomiaru głównego. Wstępny pomiar umożliwia również poprawne zaplanowanie dalszego pomiaru. Pakiet programów CrysAlis, który jest na wyposażeniu dyfraktometru, umożliwia obsługę oraz kontrolę w trakcie pomiaru. Pakiet ten składa się z dwóch programów: CrysAlisCCD, który pozwala sterować procedurami pomiarowymi oraz CrysAlisRED, służący do redukcji danych uzyskanych z pomiaru rentgenograficznego.

Z pomiaru dyfraktometrycznego uzyskuje się parametry komórki elementarnej oraz wartości intensywności promieni ugiętych dla badanego kryształu. Jest to zbiór danych eksperymentalnych niezbędnych do rozwiązania struktury.

Pierwszy etap polega na analizie tabeli pików w celu znalezienia trzech najkrótszych wektorów liniowo niezależnych. Wybór trzech niskokątowych refleksów i początku współrzędnych umożliwia znalezienie tych wektorów. Można w ten sposób zbudować komórkę elementarną w sieci odwrotnej. Następnie, program w tak zbudowanej przestrzeni próbuje wyświadczyć wszystkie uzyskane podczas pomiaru refleksy. Procedura ta może być powtarzana, program szuka nowych trzech wektorów i ponownie przypisuje wskaźniki refleksom, aż do momentu wyświadczenia większości z nich. Wskaźniki hkl, położenie refleksów (kąty  $\theta$ ) oraz długość fali użytego promieniowania program używa do wyznaczenia parametrów komórki elementarnej.

Trójwymiarowy kształt refleksów zostaje odtworzony, przy pomocy procedur, które określają położenie, podają wskaźniki hkl oraz mierzą zintegrowane intensywności refleksów wykorzystując udokładnioną macierz orientacji.

Kolejnym etapem jest redukcja danych, która polega na integracji i korekcji natężeń otrzymanych z pomiaru refleksów. Proces ten może potrwać nawet kilkadziesiąt minut, uzależnione jest to od mocy komputera oraz od liczby uzyskanych refleksów. Po redukcji danych otrzymujemy plik \*.hkl, który zawiera wskaźniki refleksów, kwadraty czynników struktury ( $F^2$ ) oraz odchylenie standardowe dla  $F^2$ . Kolejny plik (\*.sum), otrzymany na tym etapie, zawiera techniczne informacje dotyczące pomiaru, szczegóły procesu redukcji danych a także parametry komórki elementarnej. Posiadając oba pliki można przejść do dalszych obliczeń, które doprowadzą nas do rozwiązania struktury.

W celu rozwiązania problemu fazowego jak i dalszego udokładnienia struktury można skorzystać z programów opracowanych przez Georgea Sheldricka. Analizy wygaszeń systematycznych oraz wyboru grupy przestrzennej dokonuje się za pomocą programu XPREP (lub GRAL – program z pakietu CrysAlis). W przypadku związków małowartościowych, do rozwiązania problemu fazowego metodami bezpośrednimi lub metodą Pattersona stosuje się program SHELXS. W kolejnym etapie następuje udokładnienie struktury przy pomocy programu

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

SHELXL. Pracując z biomakromolekułami używa się XDS, MOSFLM oraz programy z pakietu CCP4 (SHELXC, SHELXD, SHELXE, PHASER, REFMAC). Do analizy otrzymanych wyników w sposób graficzny oraz do przygotowania rysunków związków małowcząsteczkowych stosować można program XP, ORTEP, MERCURY. W przypadku makrocząsteczek program COOT, PyMol oraz RasMol.

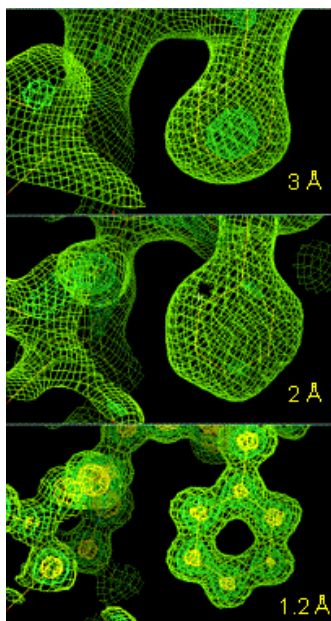
### **Ogólnie proces udokładnienia struktury składa się z następujących etapów:**

- zbudowanie modelu struktury ze wszystkimi atomami niewodorowymi,
- udokładnienie struktury z izotropowymi czynnikiemami temperaturowymi,
- udokładnienie położenia atomów niewodorowych z zastosowaniem anizotropowych czynników temperaturowych oraz lokalizacja atomów wodoru z różnicowej mapy gęstości elektronowej,
- udokładnienie wszystkich atomów z różnymi czynnikiemami temperaturowymi dla poszczególnych typów atomów (anizotropowymi dla atomów niewodorowych i izotropowymi dla atomów wodoru). Zazwyczaj nie udokładnia się położenia i czynników temperaturowych atomów wodoru. Stosowane są więzy nałożone na geometrię grup zawierających atomy wodoru, zaś wartości ich czynników temperaturowych wiąże się z czynnikami temperaturowymi atomów, z którymi są związane.

### **Wykład 6: Walidacja i deponowanie struktury**

Ważnym aspektem podczas udokładniania oraz na koniec procesu określania struktury jest ocena otrzymanego modelu. Uwagę należy zwrócić na wartości trzech parametrów, które umożliwiają ocenę „jakości” struktury. Rozdzielczość  $d$ , wyrażana w Å, określa minimalny odstęp między płaszczyznami sieciowymi kryształu poddanego promieniowaniu rentgenograficznemu. Można powiedzieć, że rozdzielczość obrazu dyfrakcyjnego świadczy o detalach jakie można rozróżnić na mapie gęstości elektronowej. Im większa wartość rozdzielczości tym bardziej dokładną mapę gęstości elektronowej można uzyskać. Rozdzielczość większa od 2Å jest rozdzielczością średnią, natomiast przy  $d$  mniejszym od 2Å mówimy o wysokiej rozdzielczości. O rozdzielczości atomowej mówimy gdy  $d$  jest poniżej 1.2Å, w tym przypadku na mapie gęstości elektronowej powinny być dobrze widoczne poszczególne atomy niewodorowe (Rysunek 6.1).

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki



**Rysunek 6.1** Mapa gęstości elektronowej przy rozdzielczości 1.2 Å, 2 Å i 3 Å.

Drugim parametrem stosowanym do oceny jakości otrzymanego modelu jest wskaźnik rozbieżności R, który wyrażany jest zazwyczaj w %. Jest „błędem względnym” wyliczonych wartości czynników struktury (Fc) w stosunku do wartości rzeczywistych (Fo).

$$R = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$$

Poprawnie udokładniona struktura związku małowcząsteczkowego powinna mieć R poniżej 10%, natomiast w przypadku struktur białkowych poniżej 20%.

Dodatkowo w trakcie udokładniania struktur białkowych kontroluje się wielkość parametru  $R_{free}$ . Jest to wartość wyliczona na podstawie około 1000 refleksów, które zostały wykluczone z udokładniania. Parametr ten jest dokładnie tak samo liczony jak wskaźnik R. Wartość  $R_{free}$  powinna być zbliżona do wartości parametru R.

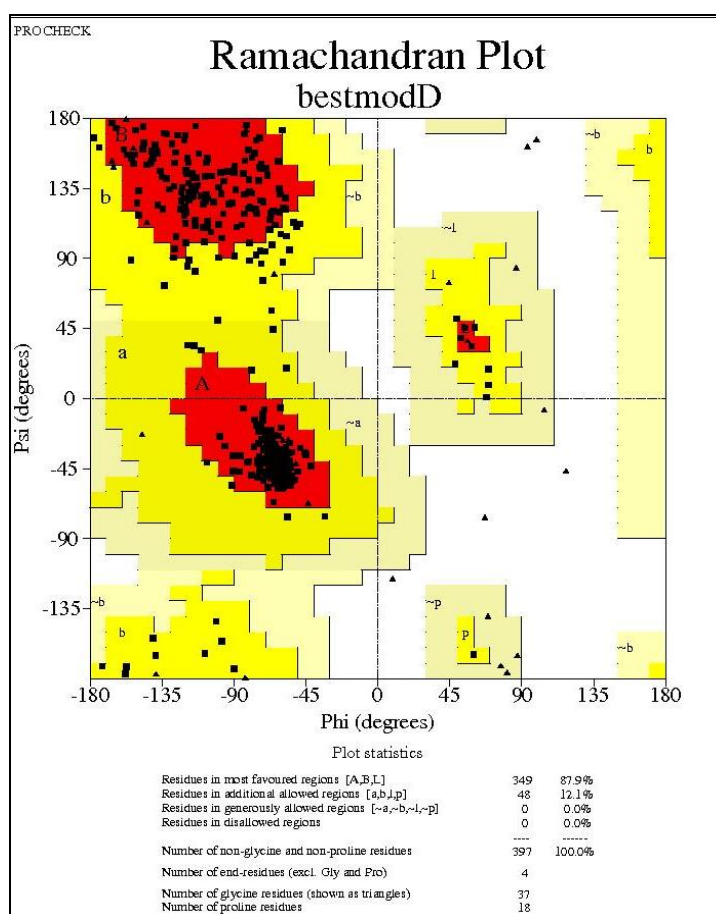
Ważnym zagadnieniem jest również sprawdzenie geometrycznej poprawności struktury. Geometria walencyjna (długości wiązań oraz wartości kątów) powinny być typowe dla struktur białkowych. Warto także zwrócić uwagę na średnią kwadratową (RMS – Root-Mean-Square). Wartość RMSD w przypadku dobrze udokładnionych struktur powinna wynosić 0.015-0.020 Å. Zbyt duża wartość (0.03 Å) lub zbyt mała (0.001 Å) świadczy o nieprawidłowym udokładnieniu.

Omówione parametry (R,  $R_{free}$ , RMSD) nie do końca są niezależnymi wskaźnikami służącymi do oceny modelu struktury, ponieważ są zaangażowane w proces udokładniania. Analiza wykresu zależności kąta torsyjnego  $\psi$  (psi) od wartości kąta torsyjnego  $\phi$  (fi) (wykres Ramachandrana - Rysunek 6.2) pozwala na obiektywne spojrzenie na otrzymany model struktury białka. Kąt  $\phi$  określa wielkość rotacji w łańcuchu głównym przy wiązaniu peptydowym między N i C $\alpha$ . Natomiast, kąt  $\psi$

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

jest mierzony między Ca i C karbonylowym. Na wykresie zaznaczone są kolorem czerwonym i żółtym regiony o dozwolonej konformacji aminokwasów, natomiast w białych obszarach znajdują się aminokwasy o niedozwolonej konformacji.

Wykres Ramachandrana dotyczy uprzywilejowanych lub dozwolonych (i niedozwolonych) regionów konformacyjnych aminokwasów dla dowolnej struktury przestrzennej białka (wszystkie aminokwasy z wyjątkiem glicyny i proliny). Poprawnie udokładniony model białka powinien posiadać co najmniej 90% reszt aminokwasowych w dozwolonych regionach. Kąty torsyjne  $\phi$  i  $\psi$  nie są uwzględniane podczas procesu udokładniania dlatego są obiektywnym kryterium geometrii.



**Rysunek 6.2** Przykładowy wykres Ramachandrana.

W ostatnim etapie otrzymany model struktury związku małowcząsteczkowego lub biomakromolekułu jest weryfikowany przy pomocy specjalnych programów. Do oceny jakości struktur związków małowcząsteczkowych stosuje się program CIFCHECK, natomiast dla struktur białkowych użyć można programu PROCHECK.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Plik CIF (Crystallographic Information File) jest typowym plikiem tekstowym oraz graficznym zawierającym wszystkie dane krystalograficzne (informację dotyczącą pomiaru oraz rozwiązanej struktury). Plik ten złożony jest z kilku części. W pierwszej części znajdują się techniczne informacje dotyczące pomiaru, dane o zastosowanych metodach i programach stosowanych do zbierania danych, redukcji danych, sposób przeprowadzenia korekcji absorpcji. Tu również zapisane są informacje odnośnie sposobu i programu użytego do rozwiązania problemu fazowego oraz do udokładnienia struktury.

W dalszej części znajdziemy parametry komórki elementarnej, rodzaj układu krystalograficznego oraz symbol grupy przestrzennej. Kolejnym fragment zawiera dane na temat upakowania cząsteczek badanego związku w strukturze oraz opis budowy poszczególnych elementów z uwzględnieniem występującej symetrii. Znajdziemy tu również opis geometrii walencyjnej (długości wiązań, kąty walencyjne), informacje o konformacji (kąty torsyjne) oraz parametry wewnątrz- i międzycząsteczkowych oddziaływań (np. wiązań wodorowych).

Dane o strukturze przestrzennej związków małowcząsteczkowych oraz biomakromolekuł znajdują się w specjalistycznych bazach, które są ogólnodostępne w Internecie lub należy za nie zapłacić. Bazy PDB (Protein Data Bank) oraz CSD (Cambridge Structure Database) są najbardziej znanymi i najczęściej używanymi tego typu bazami.

## Protein Data Bank

Protein Data Bank (PDB) jest darmową, ogólnodostępną bazą zawierającą dane o strukturze przestrzennej biomakromolekuł. Utworzona została w Brookhaven National Laboratories. Aktualnie obsługiwana jest przez The Research Collaboratory for Structural Bioinformatics.

Obecnie baza PDB zawiera 116258 struktur 3D białek, węglowodanów, kwasów nukleinowych oraz innych makromolekuł. Struktury te zostały określone metodą krystalograficzną lub techniką NMR. W bazie można również znaleźć współrzędne atomów, przesunięcia chemiczne, metodę i warunki krystalizacji, dokładne dane dotyczące pomiarów oraz dane literaturowe.

Baza PDB oraz wszystkie zawarte w niej dane dostępne są przez stronę internetową [www.rcsb.org](http://www.rcsb.org) (link aktywny styczeń 2014).

Wszystkie struktury (rekordy) znajdujące się w bazie mają nadany alfanumeryczny kod służący do ich szybkiej identyfikacji. Na kod PDB składają się cztery symbole, utworzone z zbioru cyfr od 0 do 9 i zbioru dużych liter A-Z. Niektóre, starsze kody utworzone zostały od nazwy zwyczajowej struktury.

Cała informacja krystalograficzna znajduje się w pliku o rozszerzeniu pdb. Uwzględniając różne typy informacji, plik PDB można podzielić na kilka sektorów.

- część tytułowa: zawiera opis przeprowadzonego eksperymentu oraz opis biomakrocząsteczki,
- struktura pierwszorzędowa: określa kolejność ułożenia reszt aminokwasowych w łańcuchu makrocząsteczki,



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

- hetero gen: opisuje niestandardowe reszty, takie jak grupy prostetyczne, inhibitory, aktywatory czy jony,
- struktura drugorzędowa (alfa helisy i beta kartki): identyfikuje pozycję helisy alfa w cząsteczce, przypisuje im nazwy oraz numery; zaznaczone zostają początkowe i końcowe reszty, oraz całkowitą długość helisy; identyfikuje kartki beta w sposób analogiczny jak helisy,
- wiązania chemiczne: wypisuje wszystkie mostki disiarczkowe w białkach i polipeptydach; charakteryzuje proliny i inne białka znajdujące się w konformacji cis,
- krystalografia i transformacja współrzędnych: opisuje np. parametry komórki elementarnej, grupę przestrzenną i wartość  $Z$ ,
- współrzędne atomów: zawiera numer porządkowy modelu w przypadku struktur złożonych; prezentuje współrzędne atomów reszt standardowych i niestandardowych; anizotropowy czynnik temperaturowy,
- część podsumowująca: zawiera listę liczb linii poświęconych poszczególnym rekordom,
- END Oznacza ostatni wiersz pliku.

PDB udostępnia również program RasMol, służący do wizualizacji przestrzennej cząsteczek. Dysponuje narzędziami umożliwiającymi przedstawienie cząsteczek w różnych formach przestrzennych, jak i zaznaczanie wybranych elementów. Program ten może również analizować niespójności w zbiorze danych pliku PDB, korzysta też z metody reguł chemicznych w celu określenia niektórych miejsc cząsteczki, takich jak mostki dwusiarczkowi.

## Baza CSD

CSD (Cambridge Structural Database) jest głównym produktem CCDC (Cambridge Crystallographic Data Centre). Baza CSD zawiera informacje strukturalne i bibliograficzne pochodzące z badań kryształów organicznych, metaloorganicznych i związków kompleksowych z ligandami organicznymi, metodami dyfrakcji promieni Rentgena i neutronowych. Co rok baza uzupełniana jest o dane dla nowych struktur. Graficzną wersję bazy CSD obsługują programy:

- ConQuest - przeszukiwanie bazy,
- IsoStar - ułatwia szybkie poszukiwanie i ocenę oddziaływań międzycząsteczkowych, bez konieczności konstruowania skomplikowanych zapytań lub przeprowadzić szczegółowe analizy danych.
- Mercury - wizualizacja 3D struktur oraz danych z analizy statystycznej wykonanej w bazie CSD.
- Mogul - zapewnia precyzyjne informacje na temat preferowanych molekularnych geometrii, umożliwiając dostęp do milionów chemicznie sklasyfikowanych długości wiązań, kątów walencyjnych, kątów torsyjnych czy konformacji pierścienia, dane te pochodzą z CSD.
- PreQuest - tworzenie własnych baz danych strukturalnych.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Wyniki przeszukiwania bazy danych zapisywane są w plikach – na przykład jako \*.PDF. Każdy z plików zawiera podstawowe informacje o wyszukanych związkach – takie jak: nazwa z bazy, kod z bazy, nazwa systematyczna związku, ilość cząsteczek w części niezależnej związku, autorzy publikacji, informacje o czasopiśmie, w którym ukazały się wyniki pomiarów, rozmiar i objętość komórki elementarnej, grupa przestrzenna w której krystalizuje dany związek, współrzędne położenia poszczególnych atomów, czy też właściwości chemiczne i wykorzystanie danego związku.

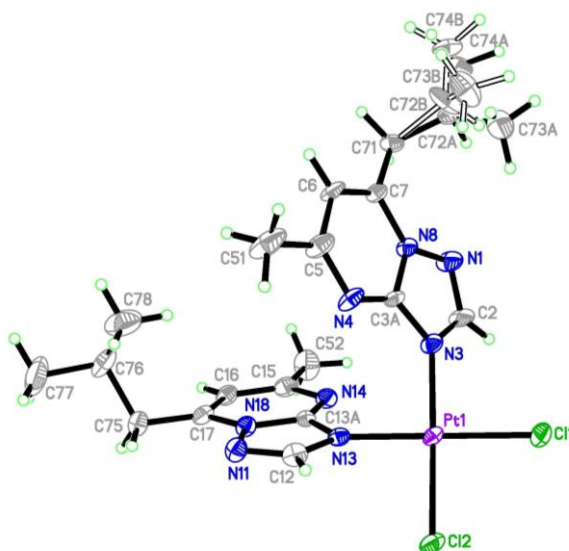
## Baza ICSD

Nieorganiczna baza danych (Inorganic Crystal Structure Database) – zawiera wszystkie struktury związków nieorganicznych, ze współrzędnymi atomów, publikowane od roku 1913. Korzystając z oprogramowania bazy ICSD, mamy możliwość wyszukiwania interesujących nas związków nieorganicznych i uzyskiwania o nich informacji zawartych w bazie.

## Wykład 7: Wizualizacja struktur białkowych oraz małowcząsteczkowych

Wizualizacja struktur 3D białek i związków małowcząsteczkowych (ligandów) odgrywa istotną rolę w analizie funkcji i aktywności złożonych biomakromolekuł. Programy do wizualizacji umożliwiają graficzne przedstawienie złożonej, wewnętrznej struktury białek.

Zazwyczaj rysunki struktur związków małowcząsteczkowych przedstawić można za pomocą jednego z programów PLUTO lub ORTEP. Rysunek 7.1 przedstawia cząsteczkę kompleksu platyny (II) przy użyciu programu ORTEP. Atomy zawsze mają kształt elipsoid, co wynika z ich drgań termicznych w sieci krystalicznej. Wielkość elipsoid powinna być proporcjonalnie zbliżona do rzeczywistej wielkości atomu.

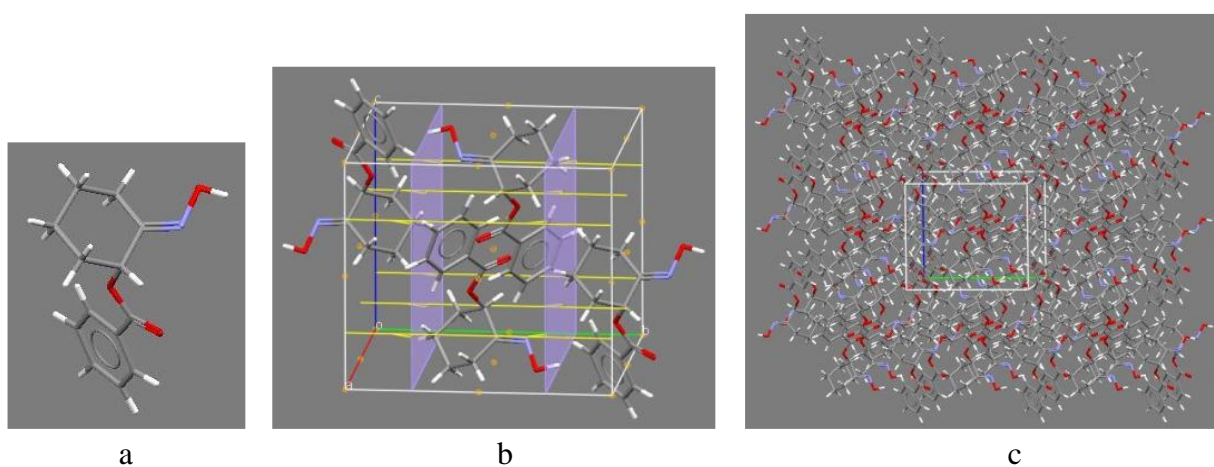


Rysunek 7.1 Typowy rysunek struktury utworzony w programie ORTEP.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

W programach do wizualizacji mają szeroką gamę różnych przydatnych funkcji. Model struktury badanego związku można oglądać w 3D, obracać go na ekranie komputera i oglądać cząsteczkę z każdej strony. Można narysować całą komórkę elementarną lub interesujący nas fragment, zmierzyć długości wiązań oraz wartości kątów walencyjnych oraz torsyjnych itd.

Również często używanym do wizualizacji ligandów jest program Mercury, którego podstawowa wersja do celów edukacyjnych dostępna jest za darmo, można ją znaleźć na stronie: <http://www.ccdc.cam.ac.uk/mercury/> (link aktywny styczeń 2014).



**Rysunek 7.2** Struktura krystaliczna przykładowego związku małowcząsteczkowego: a) pojedyncza molekula, b) komórka elementarna, c) ułożenie cząsteczek w kryształ (program Mercury).

Zbiory wejściowe używane do pracy z programem Mercury mają rozszerzenie cif, res, pdb. Pliki te zawierają pełną informację krystalograficzną, informację o położeniu atomów, geometrii walencyjnej i konformacji. Wczytując odpowiedni plik na ekranie komputera wyświetlona zostaje cząsteczka związku (Rysunek 7.2a).

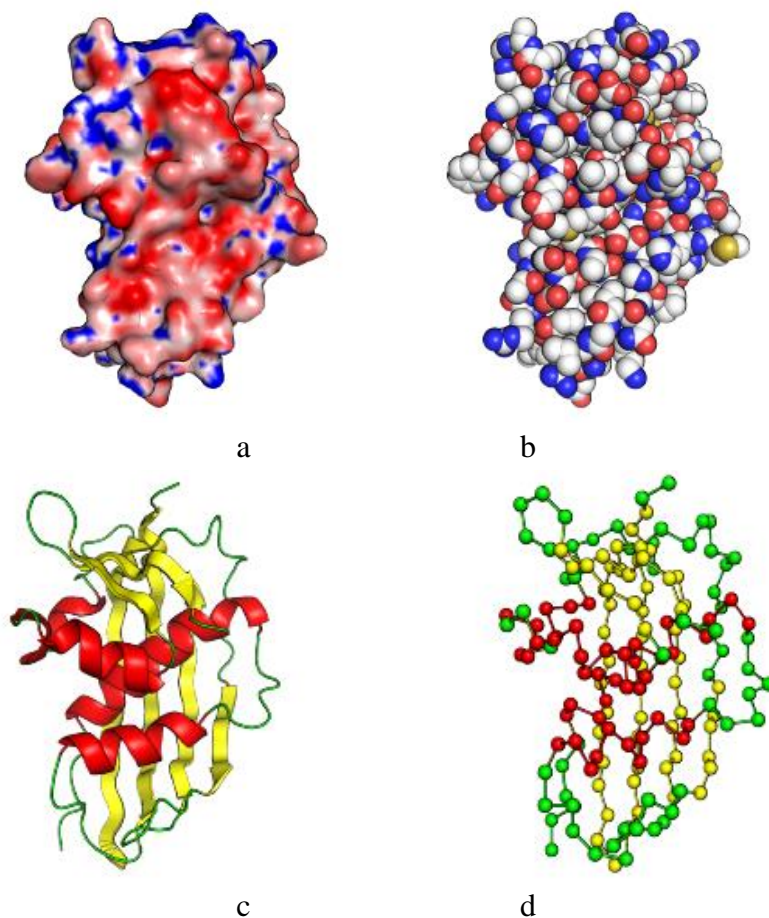
W podstawowej wersji programu można zmierzyć długość wiązań, kątów walencyjnych i torsyjnych oraz kąty międzypłaszczyznowe. Można wyświetlić krótkie kontakty międzyatomowe oraz wiązania wodorowe. Program umożliwia także analizę upakowania cząsteczek w komórce elementarnej oraz w sieci krystalicznej (Rysunek 7.2b). Możemy również rozszerzać upakowanie cząsteczek w wybranych kierunkach o określony odcinek translacyjny (Rysunek 7.2c).

Ciekawą opcją dostępną w programie jest możliwość analizowania kilku struktur jednocześnie (Multiple Structures), co ułatwia ich porównywanie. Można także nakładać cząsteczki różnych związków, co pozwala porównać ich budowę i konformację.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Pracując w programie Marcury efekty naszej pracy można zapisać w postaci „sesji” jako plik z rozszerzeniem mrx. Program umożliwia również wygenerowanie plików graficznych (jpg, tif).

Do wizualizacji złożonych cząsteczek biomakromolekuł (receptorów, kompleksów białko-ligand) można użyć różne programy np. PyMol, RasMol lub Swiss-PDBViewer. Analiza struktury przestrzennej białka przy pomocy programów graficznych dostarcza naukowcom wielu cennych informacji. Używając takich programów strukturę przestrzenną (3D) białka można przedstawić w różny sposób (Rysunek 7.3).



**Rysunek 7.3** Różne sposoby wizualizacji struktury przestrzennej białka: a) powierzchnia dostępna dla środowiska wodnego, b) model pełnoatomowy, c) schematyczny rysunek elementów struktury drugorzędowej, d) reprezentacja samych atomów Ca.

Program RasMol jest graficznym programem komputerowy, służący do wizualizacji zarówno cząsteczek białek jak i kwasów nukleinowych i małych cząsteczek. RasMol jest wartościowym programem wykorzystywanym przez biologów, biochemików oraz krystalografów. Jest programem darmowym, dostępnym na stronie: <http://www.umass.edu/microbio/rasmol/getras.htm> (link

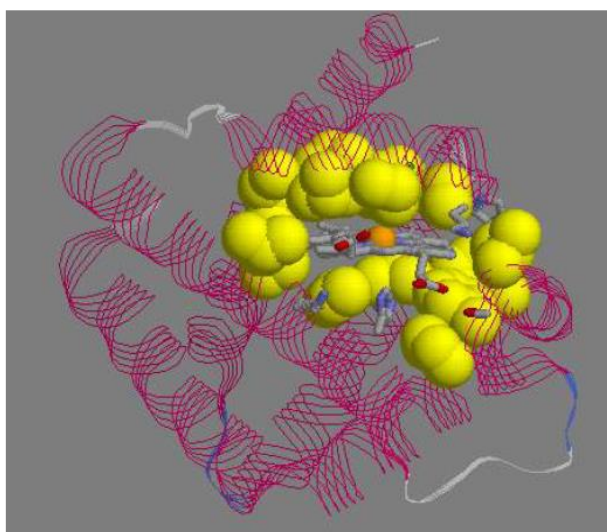
Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

aktywny: styczeń 2014). Przy użyciu tego programu możemy przedstawić cząsteczkę w różnych formach przestrzennych, w różnym ułożeniu, zaznaczyć konkretne atomy, grupy atomów, mostki disiarczkowe, wiązania wodorowe.

Omawiany program nie posiada biblioteki struktur, możemy w nim jedynie odtwarzać gotowe modele struktur, które można znaleźć w różnych darmowych bazach. Poniżej znajduje się kilka linków do baz nie tylko z danymi strukturalnymi:

- <http://www.sci.ouc.bc.ca/chem/molecule/molecule.html> (link aktywny: styczeń 2014) (baza zawierająca około 1250 związków posegregowanych tematycznie np. aldehydy, aromatyczne, leki, witaminy itp).
- <http://mc2.cchem.berkeley.edu/Smells/> (link aktywny: styczeń 2014) (baza związków zapachowych),
- <http://www.nyu.edu/pages/mathmol/library/> (link aktywny: styczeń 2014),
- <http://www.rcsb.org/pdb/> (link aktywny: styczeń 2014) (baza biomakromolekuł, Protein Data Bank).

Program PyMol jest kolejnym narzędziem używanym do wizualizacji molekuł białek, kwasów nukleinowych jak i małych cząsteczek (ligandów). Przykładowy prezentacja cząsteczki białka wykonana przy pomocy tego programu znajduje się na Rysunku 7.4. Jest to najczęściej używany program graficzny do wizualizacji biomakrocząsteczek, jego podstawowa wersja jest darmowa ([www.pymol.org](http://www.pymol.org) link aktywny: styczeń 2014). Umożliwia analizę danych strukturalnych oraz wizualizację otrzymanych modeli. PyMol posiada wiele ciekawych i przydatnych funkcji i warto się z nim zapoznać.



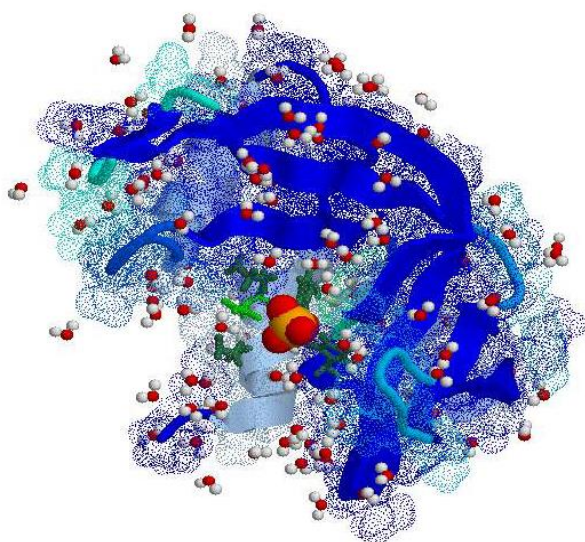
#### 1GDJ

- Tło szare
- Łańcuch główny – kolor wg struktury
- Grupa hemowa – stiks
- Atom żelaza – promień 1.2Å
- Wyróżnione **łańcuchy boczne** aminokwasów w promieniu 6.0 Å od hemu
- Tylko reszty hydrofobowe w postaci żółtych kulek

**Rysunek 7.4** Przykład wizualizacji 1GDJ w programie PyMol.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Kolejny program warty wspomnienia to Swiss-PdbViewer (<http://spdbv.vital-it.ch/> link aktywny: styczeń 2014), który pozwala na oglądanie struktur trójwymiarowych. Przykładowa wizualizacja przy pomocy tego programu znajduje się na Rysunku 7.5. Program jest łatwy w obsłudze, umożliwia też jednoczesną pracę z wieloma strukturami.



#### 5RSA

- Tło białe
- Łańcuch główny –kolor b-struktury niebieskie, helisy jasnoniebieskie
- Cząsteczki wody (DOD – ciężka woda)
- Ligand jon  $\text{PO}_4^-$  ([PO4])
- Wyróżnione **łańcuchy boczne** aminokwasów w promieniu 6.0 Å od [PO4]
- Tylko reszty polarne - GreenBlue
- Reszty dodatnie - Green

**Rysunek 7.5** Przykład wizualizacji 1GDJ w programie Swiss-PDB-Viewer.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Wykład 8: Dokowanie ligandów do receptora

Dokowanie molekularne jest modelowaniem molekularnym, które umożliwia określenie optymalnego (o najniższej energii) umiejscowienia cząsteczki liganda w strukturze receptora. Dzięki dokowaniu można przewidzieć sposoby międzycząsteczkowych oddziaływań oraz powinowactwo do wiązania w układzie ligand-białko. Technika ta odgrywa istotną rolę w procesie projektowania cząsteczek leków, jednak warunkiem koniecznym jest posiadanie strukturalnej bazy ligandów. Proces ten jest alternatywą dla eksperymentalnych metod, umożliwia wybranie z pośród dużej grupy potencjalnych leków te które będą testowane w laboratoriach.

Celem dokowania molekularnego jest znalezienie położenia i konformacji liganda w centrum aktywnym enzymu. Z reguły dokowaniu do cząsteczki receptora poddaje się kilkanaście a nawet kilkadziesiąt ligandów. Tak duża liczba związków do przetestowania wymusza zastosowania pewnego kompromisu między dokładnością, a szybkością wiązania. W ten sposób można znaleźć cząsteczki, które najsilniej oddziałują z badanym białkiem.

### Przygotowanie układu

W optymalnej sytuacji proces dokowania wymaga znajomości kompletnej struktury receptora oraz umiejscowienia i budowy centrum aktywnego. Informacja ta pozwala dokładnie sprecyzować obszar receptora, który powinien oddziaływać z cząsteczką liganda. Potrzebna jest również strukturalna baza potencjalnych ligandów.

Ważnym aspektem jest również analiza miejsca wiążącego. Powinno się zwrócić uwagę na grupy donorowe i akceptorowe dla wiązań wodorowych. Dwudodatnie jony metali często występują w centrach wiążących receptorów, jeśli zostawimy je podczas dokowania, musimy zadbać o ich poprawną parametryzację (przypisanie właściwego ładunku).

Również cząsteczki wody pełnią istotną rolę w miejscu wiążącym, mogą brać udział w „mostkowaniu” wiązań wodorowych, dlatego warto je uwzględnić jako element centrum aktywnego w trakcie dokowania. Jeśli znamy aminokwasy, które mogą tworzyć miejsce wiążące warto jest zdefiniować więzy farmakoforyczne. Pozwalają one wskazać miejsca, w którym powinien zostać umieszczony określony fragment molekuly (np. donor wiązania wodorowego). Ligandy używane do dokowania powinny być również odpowiednio przygotowane, potrzebujemy pełnoatomową strukturę oraz atomy powinny mieć przypisany właściwy ładunek.

### Procedura dokowania

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

Procedura dokowania liganda do cząsteczki receptora składa się z dwóch głównych etapów: próbkowania i oceny punktowej. Pierwszy etap: próbkowanie przestrzeni dokowania ligand-receptor polega na analizie różnych konformacji badanych ligandów w centrum aktywnym receptora. Jest to możliwe dzięki przeprowadzeniu próbkowania konformacji liganda jak i elastyczności cząsteczki białka.

Drugim etapem jest ocena punktowa, która polega na weryfikacji i punktowaniu powinowactwa do wiązania poszczególnych konformacji liganda. Pozwala to stworzyć listę ligandów w kolejności od najmniejszego do największego wyliczonego powinowactwa. Ułożenie i konformacja najwyższej ocenionego związku określa preferowany sposób wiązania w układzie ligand-białko.

## Metody dokowania

Można wyróżnić dwie metody dokowania: dokowanie oparte na próbkowaniu konformacyjnym liganda oraz próbkowaniu konformacyjnym receptora.

Procedura modelowania w oparciu o konformację liganda polega na dopasowaniu struktury liganda oraz centrum aktywnego w cząsteczce białka, algorytmy stochastyczne lub poszukiwanie systematyczne. W próbkowaniu konformacji receptora można zastosować różne algorytmy: miękkie dokowanie, relaksację molekularną, próbkowanie konformacyjne łańcuchów bocznych lub z udziałem różnych konformerów białka.

Jako funkcje oceniające stosowane są algorytmy o charakterze doświadczalnym, oparte na wiedzy, bądź wykorzystujące pola sił mechaniki molekularnej.

## Próbkowanie liganda

W najłatwiejszej metodzie dokowania opartej na próbkowaniu ligandów dopasowuje centrum aktywne w receptorze i powierzchnie molekularne dokowanych cząsteczek. Metoda ta posiada wady i zalety. Prostota i wydajność obliczeniowa jest dużą zaletą tej metody. Wadą jest sztywna konformacja badanego liganda w trakcie jego dopasowania do kształtu miejsca wiążącego.

Inna metoda polega na systematycznym poszukiwaniu. Umożliwia tworzenie wszystkich możliwych konformacji liganda, w trakcie oddziaływania z białkiem, uwzględniają jego wszystkie stopnie swobody. Można tego dokonać przez systematyczną rotację wszystkich możliwych do rotowania wiązań liganda. Trzeba jednak pamiętać, że zastosowanie takiego algorytmu generuje dużą liczbę kombinacji do przeanalizowania. Można uporać się z tym problemem nakładając w etapach początkowych więzy na niektóre wiązania w cząsteczce liganda, które są zdolne do rotacji.



Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

W metodzie systematycznego poszukiwania można zastosować również procedurę tzw. fragmentacji. Umożliwia ona podzielenie liganda na kilka sztywnych fragmentów, które są dopasowywane w różnych konformacjach do centrum aktywnego cząsteczki białka. Ligandy przyłączane są kowalencyjnie co skutkuje utworzeniem wyjątkowych ułożeń liganda.

Kolejny algorytm próbkowania wykorzystuje wcześniej przygotowane sztywne konformacje liganda.

W metodzie próbkowania ligandów można również wykorzystać algorytmy stochastyczne. Algorytm ten polega na przypadkowych zmianach w konformacyjnej, translacyjnej oraz rotacyjnej przestrzeni cząsteczek ligandów. Wprowadzone zmiany są odrzucane lub akceptowane na podstawie kryterium probabilistycznego. Algorytmy stochastyczne wykorzystywane są w wielu metodach, metodzie Monte Carlo, optymalizację wielocząsteczkową lub algorytm ewolucyjny.

## **Próbkowanie enzymu białkowego**

W białku po związaniu cząsteczki liganda może zajść dużo konformacyjnych zmian. Mogą to być tylko niewielkie zmiany w łańcuchach bocznych białka ale również zmiany obejmujące ruch większych fragmentów cząsteczek białka (domen). Białka są bardzo dużymi i elastycznymi molekułami, dlatego posiadają dużą liczbę stopni swobody. Jest to trudne do uwzględnienia w procesie dokowania.

Metoda „miękkiego dokowania” jest podstawową metodą próbkowania opartego na próbkowaniu przestrzeni konformacyjnej receptora. Metoda ta pozwala na niewielkie nakładanie się struktury liganda i białka. Jest to wydajna obliczeniowo i łatwa do wdrożenia. Zapobiega konfliktom sterycznym a jednocześnie zachowuje określone ułożenie ligandów, które może okazać się ułożeniem prawidłowym - jest to zaletą tej metody. Jednak, miękkie dokowanie może dostarczyć większej liczby niepoprawnych ułożeń liganda, które zostały zaakceptowane we wstępnym etapie.

Inną metodą jest technika próbkowania konformacyjnego łańcuchów bocznych. Szkielet cząsteczki białka pozostaje sztywny, natomiast łańcuchy boczne ulegają rotacjom. Na początku metoda ta wykorzystywała bazy rotamerów bocznych łańcuchów białkowych i korzystała z algorytmów, dzięki którym proponowane były różne kombinacje ułożenia liganda i konformerów łańcuchów bocznych.

Obecnie uwzględniana jest ciągła lub częściowa elastyczność łańcuchów bocznych. Metoda ta może bazować na algorytmie genetycznym, w którym wykorzystuje się obserwowane doświadczalnie konformacje badanego białka oraz całkowitą giętkość liganda. Technika ta wykorzystuje metodę Simulated Annealing (symulowane wyżarzanie). To rodzaj algorytmu przeszukującego przestrzeń przypadkowych zmian konformacji łańcuchów bocznych w celu

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

wyszukania najlepszego rozwiązania. Wykorzystywana jest również hipoteza, że małe rotacje kątów torsyjnych (około  $30^\circ$ ) są uprzywilejowane w stosunku do rotacji o większy kąt.

Podsumowując, próbkowanie konformacyjne łańcuchów bocznych jest niewystarczające do modelowania dużych zmian w konformacji, ponieważ nie uwzględnia elastyczności głównego szkieletu białkowego. Natomiast, dość dobrze modeluje niewielkie zmiany.

Alternatywą dla metody próbkowania konformacyjnego łańcuchów bocznych jest relaksacja molekularna. Pozwala ona na modelowanie również większych fragmentów, domen białka w przeciwieństwie do wyżej omówionej metody. Metoda ta polega na dopasowaniu sztywnego liganda do centrum aktywnego enzymu, a następnie atomy głównego szkieletu, łańcuchy boczne oraz otoczenie centrum aktywnego, w celu określenia równowagi konformacyjnej, są poddane relaksacji. Podczas relaksacji stosowana jest symulacja dynamiki molekularnej lub Monte Carlo, która umożliwia rozszerzenie próbkowania przestrzeni konformacyjnej. Można w ten sposób wyeliminować steryczne kolizje wynikające z zastosowania sztywnej cząsteczki liganda. Omawiana metoda wykazuje większe wymagania w doborze funkcji oceniających oraz jest bardziej kosztowna obliczeniowo.

Próbkowanie z udziałem wielu konformacji białka jest kolejną, wartą omówienia metodą. Pozwala ona na modelowanie elastyczności cząsteczki białka z wykorzystaniem wielu struktur w różnych konformacjach. Analizowane mogą być struktury krystalograficzne, pozyskane techniką NMR, a także uzyskane metodą symulacji dynamiki molekularnej, Monte Carlo czy Simulated Annealing. Metoda ta jest chętnie stosowana ze względu na poszerzenie przestrzeni konformacyjnej dla cząsteczkowego rozpoznania związków małowcząsteczkowych. Natomiast jej wadą jest obowiązkowe, oddzielne dokowanie dla poszczególnych konformerów białka, co jest bardzo czasochłonne. Można rozwiązać ten problem, nakładając na siebie poszczególne konformery, później są one scalane w jeden obiekt, wykorzystując go do pojedynczej symulacji.

## Funkcje oceniające

Znane są trzy różne grupy funkcji oceniających. Jako pierwsze zostaną omówione empiryczne funkcje oceniające, w których ważone doświadczalnie człony energii zostają zsumowane. Stosowane są takie wagi poszczególnych członów, które będą odtwarzały powinowactwo do wiązania ligandów ze zbioru testowego. Kolejna metoda oparta na wiedzy, zakłada, że międzycząsteczkowe oddziaływania, które często obserwowane są w kompleksach (ligand-białko) w strukturalnych bazach danych mogą sprzyjać tworzeniu tego kompleksu. Trzecia metoda wykorzystuje pola siłowe mechaniki molekularnej. Algorytm w tej metodzie opiera się na sumowaniu energii oddziaływania pomiędzy cząsteczką białka a ligandem i zmiany energii wewnętrznej liganda, wywołane naprężeniem sterycznym podczas oddziaływania z receptorem.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## Oprogramowanie

Obecnie znana jest szeroka gama możliwych do wykorzystania programów. Programy do dokowania są zarówno komercyjne jak i dostępne za darmo. Do najczęściej używanych i rekomendowanych można zaliczyć: Autodock, Autodock Vina, DOCK, DockVision, GOLD, Glide, ICM, FlexX, FRED, PSI-DOCK. Wymienione tu programy są tylko przykładami, w Internecie można znaleźć wiele innych z dokładnymi instrukcjami postępowania w trakcie dokowania.

Projekt pn. „Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych”  
realizowany w ramach Poddziałania 4.1.1 Programu Operacyjnego Kapitał Ludzki

## 9. Literatura do części II

1. Z. Bojarski, M. Gigla, K. Stróż, M. Surowiec, Krystalografia, podręcznik wspomagany komputerowo, PWN, Warszawa, 1996.
2. Z. Bojarski, E. Łągiewka, Rentgenowska analiza strukturalna, PWN, Warszawa, 1988.
3. P. Luger, Rentgenografia strukturalna monokryształów, PWN, Warszawa, 1989.
4. Protein Purification Handbook.  
[www.fm.uit.no/info/imb/amb/courses/bio360s/pdf/handbook.pdf](http://www.fm.uit.no/info/imb/amb/courses/bio360s/pdf/handbook.pdf) (link aktywny: styczeń 2014)
5. B.D. Hames, N.M. Hooper, J.D. Houghton, Krótkie wykłady. Biochemia. Wyd. Nauk. PWN. Warszawa 1999.
6. M.E. Lienqueo, A. Mahn, L. Vasquez, J.A. Asenjo, Methodology for predicting the separation of proteins by hydrophobic interaction chromatography and its application to a cell extract. *J. Chrom. A.*, 2003, 1009 189-196.
7. J.U. Bowie, Stabilizing membrane proteins. *Curr. Opin. Struc. Biol.*, 2001, 11, 397-402.
8. L. R. Eriks, J. A. Mayor, R.S. Kaplan, A strategy for identification and quantification of detergents frequently used in the purification of membrane proteins. *Anal. Bioch.*, 2003, 323, 234-241.
9. T. Ichimura, N. Ikuta, Y. Uda, T. Horigome, S. Omata, Separation of membrane proteins solubilized with a non-denaturing detergent and a high salt concentration by hydroxyapatite high-performance liquid chromatography. *Anal. Bioch.* 1995, 224, 250-255.
10. P.J. Loll, Membrane protein structural biology: the high throughput challenge. *J. Struc. Biol.*, 2003, 142 144-153.
11. J. M. Berg, L. Stryer, J. L. Tymoczko „Biochemia" Wydawnictwo Naukowe PWN 2007.
12. M. T. Bergfors, Proteine Crystallization. Techniques, strategies, and tips. A Laboratory Manual, 1999.
13. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 1993, 26, 283-291.
14. <http://www.lightsource.ca/education/synchro.php> (link aktywny: styczeń 2014)
15. <http://checkcif.iucr.org/> <http://www.cryst.chem.uu.nl/platon/> (link aktywny: styczeń 2014)
16. <http://www.wwpdb.org/documentation/format23/v2.3.html> (link aktywny: styczeń 2014)
17. CCDC: <http://www.ccdc.cam.ac.uk/> (link aktywny: styczeń 2014)
18. <http://www.ccdc.cam.ac.uk/mercury/> (link aktywny: styczeń 2014)
19. [http://matilinspnb.dbv.pl/viewpage.php?page\\_id=1096](http://matilinspnb.dbv.pl/viewpage.php?page_id=1096) (link aktywny: styczeń 2014)
20. Swiss-PDB-Viewer: <http://spdbv.vital-it.ch/> (link aktywny: styczeń 2014)