

## ZASTOSOWANIE MIKROSPEKTROSKOPII ABSORPCYJNEJ W PODCZERWIENI ORAZ MODELU REGRESJI LINIOWEJ DO ANALIZY EX VIVO STRUKTUR DRUGORZĘDOWYCH BIAŁEK W TKANKACH ZWIERZĘCYCH

Katarzyna Majzner<sup>1</sup>, Tomasz Wróbel<sup>1,2</sup>, Małgorzata Barańska<sup>1,2</sup>

<sup>1</sup>Zespół Obrazowania Ramanowskiego, Zakład Fizyki Chemicznej, Wydział Chemii Uniwersytetu Jagiellońskiego,  
ul. R. Ingardena 3, 30-060 Kraków, majzner@chemia.uj.edu.pl

<sup>2</sup>Jagiellońskie Centrum Rozwoju Leków (JCET), ul. Bobrzyńskiego 14, 30-348 Kraków

**Streszczenie:** Mikrospektroskopia FT-IR w połączeniu z chemometrią jest jedną z najważniejszych współczesnych technik analitycznych wykorzystywanych w badaniach materiałów biologicznych, takich jak tkanki czy komórki. Pozwala ona na identyfikację oraz badanie przestrzennej dystrybucji składników biochemicznych w badanym materiale, zapewniając jednocześnie wysoki poziom selektywności i rozdzielczości. Chemometria, wykorzystująca metody komputerowe, statystyczne oraz matematyczne w analizowaniu danych chemicznych, stanowi potężne narzędzie w badaniu tkanek zwierzęcych w celu obserwowania zmian chorobowych. W niniejszej pracy przedstawiono zastosowanie modelu regresji liniowej oraz spektroskopii absorpcyjnej w podczerwieni (FT-IR) do analizy zmian w profilu biochemicznym tkanki wywołanych przez chorobę cukrzycową.

**Słowa kluczowe:** Mikrospektroskopia FT-IR, chemometria, regresja liniowa, struktura drugorzędowa białek, cukrzyca

### Application of infrared absorption microspectroscopy and linear regression model to ex vivo analysis of secondary structures of proteins in animal tissues

**Abstrakt:** *Abstrakt: FT-IR microspectroscopy in combination with chemometrics is one of the most important modern analytical techniques used in studying biological materials, such as tissues or cells. It allows for identification and studying the spatial distribution of biochemical components in the sample while providing a high level of selectivity and resolution. Chemometrics, which is based on computational, statistical and mathematical methods to analyze chemical data, is a powerful tool in the study of animal tissues and observing the lesions. This paper presents an application of a linear regression model and infrared absorption spectroscopy (FT-IR) to analyze of changes in the biochemical profile of tissue caused by diabetic disease.*

**Keywords:** *Keywords: Microspectroscopy FT-IR, chemometric, linear regression, secondary structure of proteins, diabetes*

#### 1. WSTĘP

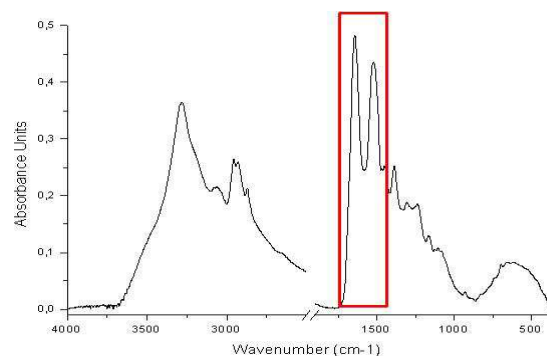
W ostatnich latach obserwuje się wzrost zainteresowania metodami analizy *ex vivo* tkanek zwierzęcych za pomocą metod spektroskopowych, mających na celu diagnostykę chorób. Zmiany chorobowe zwykle wiążą się ze zmianami struktury, budowy czy też funkcji chorego organu lub tkanki. Nic więc dziwnego, że zaczęto doświadczać poszukiwać korelacji pomiędzy zdiagnozowaną chorobą a

występującymi równoległe zmianami w profilu biochemicznym tkanek, rozumianym nie tylko jako skład poszczególnych komponentów biochemicznych, takich jak białka, tłuszcze, czy kwasy nukleinowe, ale uwzględniając również ich strukturę. Wraz z rozwojem technik spektroskopii oscylacyjnej rośnie udział fourierowskiej mikrospektroskopii absorpcyjnej w podczerwieni (FT-IR) w obrazowaniu biochemicznym tkanek. Spektroskopia absorpcyjna w zakresie średniej i bliskiej podczerwieni jest jedną z najważniejszych technik analitycznych, a jej zastosowanie do celów biomedycznych znane jest od lat.

Jedną z zalet spektroskopii absorpcyjnej w IR jest to, że praktycznie każda próbka, w dowolnej formie, może być poddana analizie. Dodatkowo, ze względu na to, że wszystkie obserwowane na widmie IR pasma pochodzą od charakterystycznych drgań dla danej grupy funkcyjnej, spektroskopia w podczerwieni pozwala na identyfikację struktur wielu związków o znaczeniu biologicznym, takich jak: białka, tłuszcze, węglowodany czy też kwasy nukleinowe. Połączenie spektroskopii absorpcyjnej w podczerwieni z mikroskopią optyczną pozwala na prowadzenie badań w mikroobszarach (z mikrometrową rozdzielczością), a tym samym na analizę biochemiczną na poziomie tkankowym, a nawet pojedynczych komórek. Dodatkowo zastosowanie metod chemometrycznych, opartych na technikach komputerowych, statystycznych oraz matematycznych w analizowaniu danych chemicznych, stanowi potężne narzędzie w analizie tkanek zwierzęcych w celu obserwowania zmian chorobowych na poziomie biochemicznym. Mikrospektroskopia FT-IR pozwala nie tylko na analizę składu biochemicznego i dystrybucji poszczególnych biochemicznych komponentów w tkance, ale również daje możliwość monitorowania zmian udziału struktur drugorzędowych białek obserwowanych w stanach chorobowych, takich jak na przykład cukrzyca czy miażdżyca. Znalezienie korelacji pomiędzy zmianami strukturalnymi białek w tkankach zwierzęcych a zmianami wywołanymi cukrzycą może stać się markerem monitorowania zmian chorobowych i określania stopnia zaawansowania choroby. Narzędziem wykorzystywanym w tym celu jest model liniowej regresji wielokrotnej oparty na pomiarach FT-IR odpowiednich wzorców białek, pozwalający na estymację dystrybucji i zawartości wybranych struktur drugorzędowych białek w tkankach zwierzęcych.

W literaturze dość dokładnie opisany został wyżej wspomniany związek na przykładzie dwóch chorób -

miażdżycy i cukrzycy. W przypadku miażdżycy jej wpływ na budowę tkanek, głównie śródbłonna, jest dość oczywisty, ponieważ obszary chorobowe bogate są w zmiany zwyrodnieniowo-wytwórcze. Ciekawym jest jednak fakt, że również znaczne różnice w profilu biochemicznym ściany naczyń może wywołać cukrzyca. W literaturze niejednokrotnie opisywana jest korelacja pomiędzy zmianami chorobowymi a zmianami profilu biochemicznego obserwowana za pomocą spektroskopii FT-IR [1].



**Rysunek 1** Przykładowe widmo FT-IR białka (hemoglobina) z zaznaczonymi pasmami amidowymi I i II.

Celem niniejszej pracy była ilościowa ocena zmian struktur drugorzędowych białek w tkance pod wpływem czynnika chorobowego. W spektroskopii absorpcyjnej w IR możemy obserwować 5 pasm białkowych, zwanych inaczej amidowymi, ponieważ pochodzą od drgań charakterystycznych wiązania amidowego – są to: amidowe I, amidowe II, amidowe III, amidowe A i amidowe B (Tabela 1, Rysunek 1).

**Tabela 1** Zakresy spektralne występowania pasm amidowych na widmie IR wraz z przypisaniem drgań dominujących ( $\nu$ -drganie rozciągające,  $\delta$ -drganie deformacyjne,  $\tau$ -drganie zginające)

Pasmo amidowe	Zakres występowania [cm <sup>-1</sup> ]	Przypisanie pasm odpowiednim drganiom	Procentowy udział drgań [%]
I	1710–1650	drganie rozciągające wiązania C=O	80% $\nu$ (C=O) (w płaszczyźnie), $\delta$ (N-H)
II	1570-1520	drganie deformacyjnych wiązania N-H	60% $\delta$ (N-H), 40% $\nu$ (C-N)
III	1400-1200	drganie rozciągające wiązania C-N	40% $\nu$ (C-N), 20% $\nu$ (CH <sub>3</sub> -C), 30% $\tau$ (N-H) (w płaszczyźnie),
A	3300-3270	drganie rozciągające wiązania N-H	$\nu$ (N-H)
B	3100-3030	drganie rozciągające wiązania C-N	$\nu$ (C-N)

## 2. METODY, APARATURA I MATERIAŁY

### 2a Metody

W czasie badań zastosowano obrazowanie za pomocą spektroskopii absorpcyjnej w podczerwieni z transformatą Fouriera w trybie transrefleksji. Badana próbka umieszczana jest na podłożu nieprzepuszczalnym dla promieniowania IR, które padając przechodzi przez próbkę i ulega odbiciu od jej powierzchni. Wiązka promieniowania, po odbiciu od powierzchni nieprzepuszczalnej, przechodzi ponownie przez próbkę i osłabiona o podwójną absorpcję dociera do detektora. Technika transrefleksji obejmuje więc refleksję na podłożu i podwójną transmisję a uzyskane widmo refleksyjno-transmisyjne nie różni się zbyt od widma transmisyjnego. Obrazowanie w podczerwieni pozwala na uzyskanie danych o charakterze spektralnym (także chemicznym) połączonych z informacją przestrzenną. W przypadku klasycznych pomiarów spektroskopowych rejestrowana intensywność jest funkcją długości fali ( $I(\lambda)$ ), z kolei w technikach obrazowania sygnał stanowi zarówno funkcję długości fali jak i położenia ( $I(\lambda, x, y)$ ), gdzie  $x$ ,  $y$  rozumiane są jako współrzędne punktu obrazowanej próbki, dla którego zostało zarejestrowane widmo IR. W ten sposób generowana jest trójwymiarowa baza danych, w której każdy punkt otrzymanego obrazu próbki reprezentuje oddzielne widmo. Na przykład użycie matrycy detektorów FPA o rozmiarze 128x128 pozwala na zarejestrowanie obrazu o rozmiarach 128x128 pikseli, gdzie każdy piksel reprezentuje oddzielny punkt próbki, dla którego zostało zarejestrowane widmo. W ten sposób, wskutek jednorazowego pomiaru techniką obrazowania, uzyskuje się 16 384 widm. System akwizycji danych 3d

### 2b Aparatura

Pomiary obrazowania wykonywane były za pomocą spektrometru Varian FT-IR 620, który wyposażony jest w: mikroskop IR 670 z obiektywem typu cassegrain 15x, przystawkę ATR z kryształem Ge, komorę do pomiarów w trybie transmisji, refleksji i ATR. Wszystkie pomiary wykonywane były z rozdzielczością  $4 \text{ cm}^{-1}$  i matrycą FPA (Focal Plane Array) 64x64.

### 2c Materiały

Podmiotem badań były standardy 30 białek (Sigma Aldrich), o zróżnicowanej strukturze drugorzędowej, na podstawie których zbudowano bazę widm absorpcyjnych w IR, która została wykorzystana do opracowania modelu liniowej regresji wielokrotnej do przewidywania udziału struktur drugorzędowych białek. Wybrane białka

reprezentują szeroki zakres zmienności struktur helikalnych i beta-kartek oraz różnorodne domeny sfałdowań. Zebrana grupa białek podzielona została na 4 klasy białek, zgodnie z klasyfikacją w bazie danych SCOP (ang. *Structural Classification of Proteins*), gdzie białka grupowane są w oparciu o sposób fałdowania się łańcucha. Spośród 11 wyróżnianych w bazie SCOP klas, wybrano klasy charakteryzujące się następującymi strukturami drugorzędowymi: alfa (12 białek), beta (13 białek), alfa + beta - zbudowane głównie z antyrównoległych beta-arkuszy występujących w posegregowanych jednostkach strukturalnych alfa i beta (5 białek), alfa/beta - zbudowane głównie z równoległych beta-arkuszy w jednostkach strukturalnych beta-alfa-beta (4 białka).

Białka mierzone były w postaci filmów z roztworów wodnych, w których buforem był HEPES, czyli kwas hydroksyetylodietylenodiaminoetanolosulfonowy ( $\text{C}_8\text{H}_{18}\text{N}_2\text{O}_4\text{S}$ ). Białka rozpuszczono w HEPES o stężeniu 2mM (masa molowa 238.3 g/mol) i pH ustalonym za pomocą 1M wodorotlenku sodu (NaOH) do wartości ok. 7,2. Filmy przygotowywano poprzez nanoszenie kropli roztworu białka na szkiełko o właściwościach refleksyjnych MirrIR (Kevley Technologies, Ohio, USA) i pozostawiano następnie do odparowania.

Badane metodą spektroskopii absorpcyjnej tkanki pochodzą z dwóch zwierzęcych modeli cukrzycy: mysz model cukrzycy typu 2, czyli mysz *db/db* oraz szczurzy, streptozocynowy model cukrzycy typu 1. Oznaczenie „*db*” oznacza mutacje w receptorze leptyny, natomiast „*db/db*” oznacza homozygotę dla tej mutacji. U wszystkich myszy *db/db* występuje otyłość, która wtórnie wywołuje cukrzycę typu 2. Cukrzycę streptozocynową (STZ) osiągnano poprzez dootrzewnowe podawanie zwierzętom antybiotyku, streptozotocyny, pozyskiwanego ze szczepu *Streptomyces griseus* (70 mg/kg ciała zwierzęcia). Antybiotyk ten jest inhibitorem procesu replikacji DNA i wykazuje szczególnie powinowactwo do DNA komórek  $\beta$  trzustki, powodując degradację matrycy DNA. Wywoływana u badanych szczurów cukrzyca, objawami i przebiegiem choroby odpowiada cukrzycy typu I. Grupie kontrolnej podawano dootrzewnowo roztwór soli fizjologicznej.

Preparatyka przygotowania tkanek polega na pobraniu organu (tu: aorta piersiowa), zanurzenie w preparacie pozwalającym na cięcie zamrożonych tkanek (ang. Optimum Cutting Tool, OCT), pozostawienie na kilkanaście godzin w temp. - 80°C, a następnie cięcie przekrojów o grubości 10  $\mu\text{m}$  za pomocą mikrotomu. Następnie tkanki nanoszone były na szkiełko o właściwościach refleksyjnych MirrIR i mierzone techniką transrefleksji. Parametry pomiarów techniką obrazowania w trybie transrefleksji:

zakres spektralny - 850-3800  $\text{cm}^{-1}$ , rozdzielczość - 4  $\text{cm}^{-1}$ , czas integracji - 0,039 ms (optymalizowany przy każdym pomiarze), detektor, rozmiar matrycy - FPA-Lancer-MCT, 64 x 64 pikseli, temperatura detektora - 76 - 80 K, częstotliwość zbierania - 5 kHz lub 2,5 kHz, liczba skanów próbki - 64, liczba skanów tła - 128.

### 3. ANALIZA I PRZETWARZANIE DANYCH

Mimo ogromnych możliwości jakie daje spektroskopia FT-IR, wnioskowanie na temat zmian w profilu biochemicznym, czy też na temat korelacji ze zmianami chorobowymi byłoby niemożliwe bez zastosowania technik eksploracji danych. W celu analizy wyników, uzyskanych techniką obrazowania w IR, wykorzystano program CytoSpec v.1.4.03 (<http://www.cytospec.com/>) - dedykowany m.in. do obrazowania hiperspektralnego. Program ten pozwala na poprawę jakości widm poprzez zastosowanie opcji wstępnego przetwarzania (np. normalizacja, wygładzanie, korekta linii bazowej, interpolacja, druga pochodna), jak i podstawową analizę chemometryczną danych (PCA, HCA, KMC, FCM, ANN). W niniejszej pracy niezwykle użyteczna okazała się hierarchiczna analiza skupień (HCA) oraz model liniowej regresji wielokrotnej, w oparciu o który wnioskowano na temat związku pomiędzy zmianami w profilu biochemicznym tkanek a zmianami chorobowymi.

Wyniki obrazowania białek oraz tkanek poddane zostały wstępnemu przetwarzaniu, polegającym m. in. na wybraniu zakresu spektralnego obejmującego pasma amidowe I i II, który najlepiej korelują ze strukturą drugorzędową białek (1450 - 1750  $\text{cm}^{-1}$ ), czy też przeprowadzeniu tzw. testu jakości (ang. Quality Test, QT), w którym odrzucane były piksele reprezentujące wartości spoza wyznaczonego zakresu, czyli pozbawione informacji na temat badanej próbki (wartości definiowane arbitralnie). HCA, pozwalająca na uzyskanie widm średnich, wykonano w oparciu o drugie pochodne widm transrefleksyjnych. W celu uzyskania drugich pochodnych wykorzystano algorytm Savitzky'ego i Golaya z zastosowaniem 9-punktów wygładzania. W analizie HCA jako metodę aglomeracji wybrano algorytm Warda'a, natomiast jako miarę odległości (rjk) wybrano metodę D-values, której podstawą jest tzw. odległość Pearsona (definiowaną poprzez współczynnik Pearsona, stanowiący miarę korelacji dwóch wektorów):

$$r_{jk} = \frac{(\sum_{i=1}^n x_{ji} \cdot x_{ki}) - n \cdot \bar{x}_j \cdot \bar{x}_k}{\sqrt{(\sum_{i=1}^n x_{ji}^2 - n \cdot \bar{x}_j^2) \cdot (\sum_{i=1}^n x_{ki}^2 - n \cdot \bar{x}_k^2)}}$$

Rezultatem analizy HCA jest wyodrębnienie wybranej liczby klas spośród wszystkich pikseli otrzymanego obrazu. Dla każdej klasy obliczane jest średnie widmo z wszystkich widm przypisanych do danej klasy. Dzięki zastosowaniu HCA udało się uzyskać widma białek, które następnie stworzyły bibliotekę danych niezbędną do budowy modelu liniowej regresji wielokrotnej. Do dalszej analizy, prowadzącej do otrzymania równań modelu regresji, wykorzystano środowisko programu MatLab 10 (The MathWorks, Natick, MA).

### 4. MODEL LINIOWEJ REGRESJI WIELOKROTNEJ

Zastosowana w niniejszej pracy metoda opiera się na korzystaniu z bazy danych zawierających widma białek o znanej strukturze drugorzędowej, co oznacza, że dla każdego białka niezbędne jest dokładne zdefiniowanie tej struktury. Każde białko wnosi do budowanego modelu unikalną informację na temat struktury, a dokładne wartości udziału struktur drugorzędowych dla każdego z nich uzyskano za pomocą algorytmu DSSP zaimplementowanego w bazie PDB (Protein Data Bank) [2, 3]. Struktury drugorzędowe wyliczone algorytmem DSSP dzielą się na 8 frakcji, zakodowanych jako struktury: H ( $\alpha$ -helisa), B ( $\beta$ -bridge), E ( $\beta$ -kartka), G ( $3_{10}$ helix), I (5 helix ( $\pi$  helisa)), T (zwroty stabilizowane wiązaniami wodorowymi), S (zagięcie, struktura „bend”), NA (struktura nieokreślona). Fragment bazy danych z przypisanym udziałem struktur drugorzędowych dla każdego białka przedstawiono w Tabeli 2.

Celem niniejszej pracy jest określenie znaczenia każdej liczby falowej analizowanego widma absorpcyjnego w IR w skutecznym przewidywaniu danej struktury drugorzędowej. Ideę tworzenia modelu służącego do predykcji zawartości struktur drugorzędowych białek można przedstawić w formie iloczynu macierzy:

$$\begin{bmatrix} 1 & A_{1,w_1} & \dots & A_{1,w_k} & A_{1,w_n} \\ 1 & A_{2,w_1} & \dots & A_{2,w_k} & A_{2,w_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & A_{34,w_1} & \dots & A_{34,w_k} & A_{34,w_n} \\ 1 & A_{35,w_1} & \dots & A_{35,w_k} & A_{35,w_n} \end{bmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix} = \begin{pmatrix} C_{SecStruct\ 1} \\ C_{SecStruct\ 2} \\ C_{SecStruct\ 3} \\ \dots \\ C_{SecStruct\ 34} \\ C_{SecStruct\ 35} \end{pmatrix}$$

Do budowy modelu opisującego średnie zawartości struktur drugorzędowych białek użyto wartości absorbancji (A) przy wybranych wartościach liczb falowych (w). Proponowany

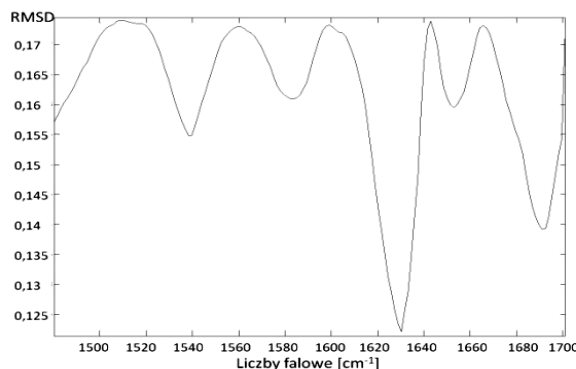
model jest liniowy dla zakresu absorpcji  $A_{j,wi}$ , gdzie  $A$  to absorpcja przeskalowanych widm,  $j$  oznacza numer widma (od 1 do 35) a  $w_i$  liczbę falową. Model obejmuje stałą  $a_1$  oraz współczynnik proporcjonalności  $a_2$ . Dla każdej liczby falowej  $w_i$ , zawartość struktury drugorzędowej białek  $f_{SecStruct j}$  jest związana z absorpcją  $A_{j, wi}$  poprzez zależność:

$$C_{SecStruct j} = a_1 + a_2 * A_{j, w_i} + \dots + a_n * A_{j, w_k}$$

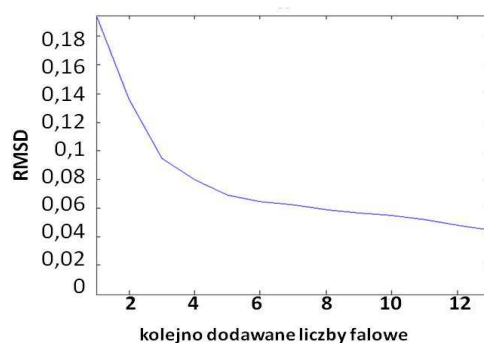
Ponieważ wielkości występujące w powyższej macierzy są niewiadomymi, które w procesie budowy modelu należy ustalić, dlatego też konieczna była znajomość zawartości struktur drugorzędowych dla każdego białka, czyli  $f_{SecStructure j}$  (Tabela 2). Wyboru kolejnych liczb falowych  $w_k$  dokonano w oparciu o najniższą wartość pierwiastka kwadratowego ze średniej arytmetycznej kwadratów odchyień (ang. root mean square deviation, RMSD; rysunek 3). RMSD jest miarą zgodności i jest najczęściej stosowanym miernikiem różnic pomiędzy wartościami przewidywanymi przez model (lub estymator) a wartościami rzeczywistymi. Wartość RMSD estymatora  $\hat{\theta}$ , w stosunku do szacowanego parametru  $\theta$ , definiowana jest jako pierwiastek kwadratowy błędu średniokwadratowego:

$$RMSD(\hat{\theta}) = \sqrt{MSE(\hat{\theta})} = \sqrt{E((\hat{\theta} - \theta)^2)}$$

gdzie MSE – błąd średniokwadratowy. Następnie szukano wartości kolejnych liczb falowych, dla których udział absorpcji (obok absorpcji przy  $w_i$ ) w całym modelu dla danej struktury drugorzędowej pozwoli obliczyć zawartość tej struktury w sposób jak najbliższy wartości rzeczywistej, czyli wartość RMSD będzie jak najniższa. Postępowanie to powtarzano do momentu uzyskania satysfakcjonującego poziomu błędu predykcji, równocześnie wybierając najmniejszą liczbę zmiennych. Wykres zależności wartości RMSD dla kolejnych liczb falowych w zakresie 1710 – 1480  $cm^{-1}$  przedstawia rysunek 3, natomiast rysunek 4 pokazuje zmianę wartości RMSD w zależności od ilości liczb falowych uwzględnianych w modelu (od 1 do 13).



**Rysunek 2** Zależność RMSD od wartości liczby falowej dla jednej zmiennej (1630  $cm^{-1}$ ) w przewidywaniu udziału struktury  $\beta$ -kartki



**Rysunek 3** Wykres zależności wartości RMSD od kolejno dodawanych do modelu wartości absorpcji przy najlepiej dobranych liczbach falowych z zakresu 1710–1480  $cm^{-1}$  (struktura  $\alpha$ -helisy)

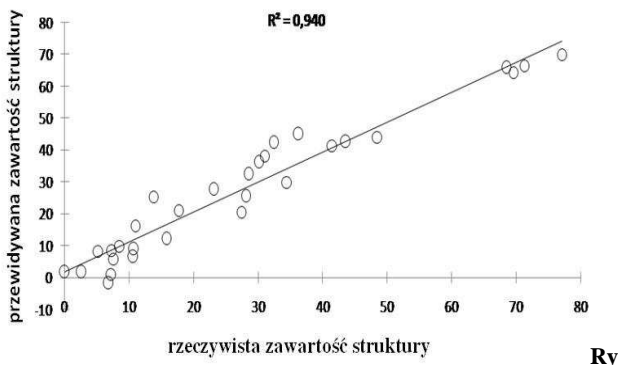
W oparciu o wykresy zależności wartości RMSD od kolejno dodawanych do modelu liczb falowych (rysunek 3) stworzono modele liniowej regresji wielokrotnej dla każdej z 7 (struktura I nie była liczona ze względu na niską zawartość w bazie – tylko 2 białka posiadały fragment o tej strukturze) struktur drugorzędowych białek zawierające optymalną liczbę zmiennych. Optymalna liczba była wybierana w sytuacji, kiedy dodanie do modelu kolejnej zmiennej nie przynosiło istotnych zmian w wartości RMSD modelu. Przykładowe równanie modelu liniowej regresji wielokrotnej, wraz z wartością RMSD, zastosowane do obliczenia zawartości alfa helisy, przedstawiono poniżej:

$$\% H = (-0,03818 * A^{1710} - 1,742 * A^{1637} - 0,8411 * A^{1547} + 8,600 * A^{1504} + 0,5433 * A^{1550} - 6,224 * A^{1599} + 0,3172 * A^{1529} - 2,290 * A^{1630} + 0,6299 * A^{1699} + 9,334 * 10^{-15}) * 100\%$$

Wyniki analizy *ex vivo* tkanek cukrzycowych oraz tkanek kontrolnych, będących referencją zmian, przedstawiono na rysunku 5 w postaci map dystrybucji każdej ze struktur. Każdy piksel obrazu, reprezentujący pojedyncze, uśrednione z 64 skanów widmo, został przeliczony za pomocą równań modelu na zawartość każdej z analizowanych struktur drugorzędowych białek. Jak widać na Rysunku 5 pod wpływem zmian cukrzycowych obserwuje się znaczne zmiany w zawartości struktur drugorzędowych białek w postaci spadku zawartości struktury  $\alpha$ -helisy i wzrost zawartości struktury  $\beta$ -kardki. Dodatkowo maleje zawartość struktur T ( $\beta$ -turn) i B ( $\beta$ -bridge) oraz wzrasta zawartość struktur B (*bend* – zagięcie) oraz NA (nieokreślonej). Wskazaną tendencję zmian obserwuje się dla obu zwierzęcych modeli cukrzycy (myszy *db/db* i szczury STZ).

Otrzymane modele dają satysfakcjonujące przewidywania zawartości sześciu liczonych struktur drugorzędowych, czego dowodem jest przedstawiony na rysunku 4 przykładowy wykres regresji liniowej dla struktury  $\alpha$ -helisy. Przedstawia on zależność pomiędzy przewidywaną i rzeczywistą zawartością struktury drugorzędowej w bazie wraz z wartością współczynnika korelacji  $R^2$  wynoszącym 0,940.

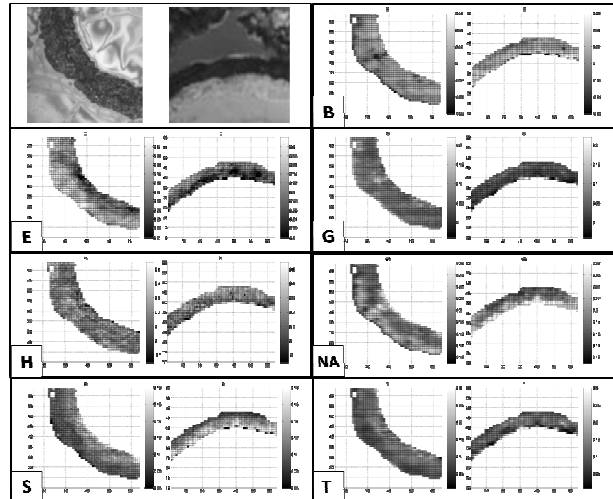
Następnie otrzymane modele liniowej regresji wielokrotnej zostały zastosowane do obliczenia zawartości struktur drugorzędowych białek w tkankach zwierzęcych.



**Ry sunek 4** Wykres linii regresji dla struktury  $\alpha$ -helisy w oparciu o przewidywaną i rzeczywistą zawartość struktury w bazie danych wraz z wartością współczynnika determinacji ( $R^2=0,940$ )

Obserwowana tendencja zmian w profilu białkowym pomiędzy tkanką cukrzycową a kontrolną jest konsekwencją zmian zachodzących w tkance w czasie procesu chorobowego. Towarzyszący cukrzycy silny stres

oksydacyjny powoduje m.in. glikację białek, która może być jedną z przyczyn zmian w strukturze drugorzędowej.



**Rysunek 5** Przekrój poprzeczny i mapy dystrybucji analizowanych struktur drugorzędowych białek otrzymane za pomocą modelu dla aorty piersiowej szczura STZ (po prawej stronie każdego zestawienia tkanka kontrolna, po lewej – cukrzycowa).

## 5. PODSUMOWANIE

Wykorzystanie mikrospektroskopii FT-IR do badań *ex vivo* struktur drugorzędowych białek oferuje nowatorskie podejście do badania złożonych procesów zachodzących w tkance. Zastosowany w niniejszej pracy model liniowej regresji wielokrotnej, a w konsekwencji obserwacja zmian profilu białkowego w tkankach cukrzycowych wykazały zmiany w zawartości struktur  $\alpha$ -helisy (spadek) oraz  $\beta$ -kardki (wzrost) w badanych tkankach cukrzycowych. Tego typu analiza może być aktualnie wykonana jedynie na tkankach pobranych post mortem, natomiast trwają prace nad próbnikami pozwalającymi zbierać widma FT-IR endoskopowo.

### Podziękowania

Badania zostały wykonane w ramach projektu pn. *Interdyscyplinarne Studia Doktoranckie „Nauki molekularne dla medycyny”* współfinansowanego ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego – Program Operacyjny Kapitał Ludzki 2007-2013. Praca współfinansowana ze środków funduszy Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego (grant koordynowany przez JCET-UJ, POIG.01.01.02-00-069/09).

## **Literatura**

1. Toryan N. et al., *Biochem. J.* (2006) 397, 427–436
2. Kabsch W., Sander C., 1983, *Biopolymers* 22:2577–2637;
2. Joosten R.P., Te Beek T.A.H., Krieger E., Hekkelman M.L., Hooft R.W.W., Schneider R., Sander C., Vriend G., (2010), *Nucleic Acids Research*, 39 (SUPPL. 1):D411-D419
3. Oberg K. A., Ruyschaert J.M., Goormaghtigh E. (2003), *Protein Science*, 12:2015–2031
4. Goormaghtigh E., Gasper R., Bénard A., Goldsztein A., Raussens V. (2009), *Biochimica et Biophysica Acta* 1794:1332–1343
5. Wrobel T.P., Malek K., Chlopicki S., Mateuszuk L., Baranska M. (2011), *Analyst*, 136(24):5247-5255
6. Bonnier F., Rubin S., Debelle L., Vente´o L., Pluot M., Baehrel B., Manfait M., Sockalingum G.D. (2008, *J. Biophoton.*, 3(1):204–214
7. Kneippa J., Millerb L.M., Joncica M., Kittelc M., Lascha P., Beekesa M., Naumanna D. (2003), *Biochimica et Biophysica Acta* 1639:152– 158.