

Imię i nazwisko autora rozprawy: Łukasz Nierzwicki
Dyscyplina naukowa: Chemia

ROZPRAWA DOKTORSKA

Tytuł rozprawy w języku polskim: Symulacje biomolekularne jako narzędzie do badania mechanizmów rozpoznania i wiązania pomiędzy białkami

Tytuł rozprawy w języku angielskim: Biomolecular simulations as a tool for studying the recognition and binding mechanisms of proteins

Promotor	Drugi promotor
<i>podpis</i>	<i>podpis</i>
dr hab. inż Jacek Czub, prof. nadzw. PG	prof. dr hab. Jarosław Marszałek

Abstract

Proteins perform a vast array of functions within living organisms, however they rarely act on their own. In most cases, they cooperate with each other, forming large molecular machineries that need a precise arrangement of their components for the proper functioning. For this reason, protein–protein interactions attract a lot of attention, as they govern the biological activity of the proteins.

Although a numerous experimental methods have been developed to examine protein–protein binding, they are mostly expensive and infrequently provide a direct access to the molecular details of the protein recognition and cooperation mechanisms. That is why there is an urgent need for methods which provide rational hypotheses that can be challenged by experiments.

In this work, I aimed to show that modern computational methods – in particular, all-atom molecular dynamics simulations – are capable of predicting both the precise structure of protein–protein complexes and the molecular driving forces that govern the recognition and binding of proteins. To explore the scope of these methods, I have examined three distinct systems: i) formation of a binary complex of the Hsp70 protein with its co-chaperone, ii) stability and intradomain interactions within 4 helix bundle (4HB) domain of Zuotin, and iii) the substrate recognition and binding by γ -secretase. These systems differ in both the forces underlying the binding affinity and the native environment, as they are either water soluble (Hsp70, 4HB) or membrane embedded (γ -secretase) proteins. Obtained results clearly show that molecular dynamics is capable of both accessing the molecular details of the interaction that rationalize the experimental findings and of providing sound predictions that can further guide the experiments.

Streszczenie

Białka pełnią wielorakie funkcje w organizmach żywych, jednak rzadko kiedy działają samodzielnie. W zdecydowanej większości przypadków współpracują one między sobą tworząc makrokompleksy o precyzyjnie zorientowanych podjednostkach, co umożliwia ich prawidłowe współdziałanie. Z tej przyczyny oddziaływania międzybiałkowe wzbudzają bardzo duże zainteresowanie, jako że w istotny sposób odpowiadają za aktywność biologiczną białek.

Choć do tej pory opracowano wielorakie metody doświadczalne badania oddziaływań białko białko, to niestety bardzo często są one kosztowne oraz rzadko kiedy pozwalają na scharakteryzowanie w bezpośredni sposób molekularnych szczegółów mechanizmów rozpoznania oraz wiązania białek między sobą. Z tej przyczyny istnieje duże zapotrzebowanie na alternatywne metody, które mogłyby dalej pokierować badaniami eksperymentalnymi.

W niniejszej pracy pragnęłam pokazać, że nowoczesne metody komputerowe, ze szczególnym wyróżnieniem pełnoatomowego wariantu dynamiki molekularnej, umożliwiają zarówno precyzyjne przewidywanie struktur kompleksów białkowych, jak również scharakteryzowanie molekularnych sił napędowych leżących u podstaw tworzenia kompleksów. By zbadać zakres stosowalności tych metod, w swojej pracy badałam trzy różnorodne systemy: i) tworzenie kompleksu przez białka Hsp70 z ich partnerami pomocniczymi, ii) wewnątrzdomenowe oddziaływania w ramach domeny 4HB białka zuotyny oraz iii) rozpoznanie i wiązanie substratów przez γ -sekretagę. Systemy te różniły się między sobą zarówno co do rodzaju sił leżących u podstaw tworzonych oddziaływań, jak również środowiskiem, w którym te białka funkcjonują. Otrzymane wyniki w jasny sposób pokazują, że symulacje dynamiki molekularnej pozwalają zarówno na uzyskanie molekularnych szczegółów oddziaływań między białkami przy pomocy których można interpretować otrzymywane wyniki doświadczalne, jak również na dostarczanie racjonalnych hipotez, które mogą pomóc w projektowaniu kolejnych badań doświadczalnych.

Podziękowania

Pragnę serdecznie podziękować prof. Jackowi Czubowi, za cierpliwość, rzetelność i gotowość do pomocy o praktycznie każdej porze dnia i nocy oraz prof. Jarosławowi Marszałkowi za włożone zaangażowanie oraz możliwość pracy w multidyscyplinarnym zespole. Dziękuję również Kasi i Asi, za cierpliwe znoszenie moich całonocnych nieobecności i udzielane wsparcie, jak również Cyprianowi i Michałowi zarówno za rozmowy, uwagi i sugestie co do tekstu, jak i za samą obecność przez czas pisania.

Spis treści

1	Wprowadzenie	1
1.1	Oddziaływania białko-białko	1
1.2	Budowa miejsc rozpoznania białek	1
1.3	Typy sił napędowych tworzenia kompleksu	3
1.3.1	Oddziaływania polarne	3
1.3.2	π -stacking	3
1.3.3	Efekt hydrofobowy	4
1.3.4	Komplementarne powierzchnie kontaktu w błonie	4
1.3.5	Niedopasowanie hydrofobowe	4
1.4	Techniki doświadczalne dla badania oddziaływań białko-białko	5
1.4.1	Techniki jakościowe	5
1.4.2	Techniki ilościowe	5
1.4.3	Strukturalne	6
1.5	Metody obliczeniowe	7
1.6	Wybrane przykłady oddziaływań białko-białko	8
1.6.1	Budowa i funkcje białek rodziny HSP70	9
1.6.2	Budowa i funkcje białka zuotyny	12
1.6.3	Budowa i funkcje γ -sekreazy	13
2	Cele oraz zakres pracy	18
3	Metody obliczeniowe	20
3.1	Elementy mechaniki statystycznej	20
3.2	Mechanika kwantowa a pola siłowe	23
3.3	Pole siłowe	25
3.4	Metody wzmocnionego próbkowania	27
3.4.1	Metody replik temperaturowych	27
3.4.2	Metadynamika	29
3.4.3	Umbrella sampling	30
3.5	Procedura symulacyjna	31
3.5.1	Procedury symulacyjne dla systemów zawierających białka HSP	32
	Relaksacja struktury modelu homologicznego białka Ssq1	32
	Spontaniczne wiązanie białka Jac1 do białka Ssq1	32
	Symulacje umbrella sampling dla asocjacji białka Jac1 do białka Ssq1	33
	Dokowanie molekularne i równowagowe symulacje kompleksu Jac1-Ssq1	33
	Relaksacja konformacyjna pętli białka Jac1	34
	Symulacje 2D metadynamiki dla oddziaływania Arg207/167 z jej natywnymi kontaktami w białku Ssq1/DnaK w obecności bądź pod nieobecność związanego białka JDP	34



Symulacje 1D metadynamiki dla zmiany partnerów wiązania Arg207 w białku Ssq1 oraz Arg167 w DnaK z natywnych na resztę Asp motywu HPD białek JDP	35
Symulacje dynamiki sterowanej dla rozdysocjowania domen NBD i NBD- β białek Hsp70 w stanie wolnym i w kompleksie z białkiem JDP	35
Symulacje umbrella sampling dla rozdysocjowania domen NBD i NBD- β białek Hsp70 w stanie wolnym i w kompleksie z białkiem JDP	36
Analiza głównych składowych dla zmian w obrębie NBD białka DnaK pod wpływem wymiany nukleotydu bądź związania domeny J białka DnaJ	36
3.5.2 Budowa symulowanych układów dla systemów zawierających domenę 4HB	37
Równowagowe symulacje domeny 4HB oraz wyznaczenie wkładów do rdzenia hydrofobowego domeny	37
Rozwijanie domeny 4HB przy pomocy symulacji dynamiki sterowanej	38
Symulacje replik temperaturowych domeny 4HB <i>S. cerevisiae</i>	38
Wyznaczanie energii swobodnej wypętlenia C-końca domeny 4HB <i>S. cerevisiae</i>	38
3.5.3 Budowa symulowanych układów dla systemów zawierających γ -sekretazę	39
Symulacje dynamiki sterowanej dla podejścia substratu do wstępnego miejsca wiązania γ -sekretazy	39
Symulacje umbrella sampling z wymianą replik dla podejścia substratu w pobliże centrum katalitycznego	39
4 Wyniki i dyskusja	41
4.1 Oddziaływanie białek HSP70 z białkami pomocniczymi JDP	41
4.1.1 Tworzenie kompleksu Jac1-Ssq1 jest napędzane przez oddziaływanie elektrostatyczne	41
4.1.2 Kompleks Jac1-Ssq1 wykazuje jeden główny sposób wiązania	43
4.1.3 Dopasowanie pętli HPD Jac1 do Ssq1 zachodzi zgodnie z mechanizmem selekcji konformacyjnej	46
4.1.4 Pętla HPD białka JDP osłabia oddziaływanie pomiędzy domenami NBD i SBD- β poprzez oderwanie kluczowej argininy	47
4.1.5 Związanie domeny J do HSP70 powoduje wstępne rozdysocjowanie domen NBD i SBD- β	49
4.1.6 Steryczne odpychanie między helisą 3 domeny J a SBD skutkuje wstępnym rozdysocjowaniem domen SBD i NBD białek Hsp70	53
4.2 Zmiany strukturalne domeny 4HB przysposabiające do aktywacji genów indukujących oporność wielolekową u drożdży <i>Saccharomyces cerevisiae</i>	56
4.2.1 W procesie ewolucji <i>Saccharomyces cerevisiae</i> znacząco wzrasta hydrofobowość C-terminalnego końca domeny 4HB	56
4.2.2 W procesie ewolucji związanie C-terminalnego końca do reszty domeny 4HB zaczęło odpowiadać za stabilizację tej domeny	58

4.2.3	Zmiany konformacyjne zachodzące w obrębie domeny 4HB są związane z wypętlaniem C-terminalnego fragmentu domeny na zewnątrz	60
4.3	Rozpoznawanie i wiązanie substratów przez γ -sekretazę	61
4.3.1	Cholesterol odgrywa istotną rolę w wiązaniu substratów przez γ -sekretazę	61
4.3.2	Błona ulega lokalnemu ścinieniu w bezpośrednim sąsiedztwie wstępnego miejsca wiązania substratu γ -sekretaży	62
4.3.3	Efekt hydrofobowy stanowi siłę napędową leżącą u podstaw asocjacji substratu do wstępnego miejsca wiązania γ -sekretaży	63
5	Wnioski	66

Wykaz skrótów

4HB	domena białka zuotyny (ang. four helix bundle)
Ala	Alanina
ATP	Adenozynotryfosforan
Arg	Arginina
Asn	Asparagina
Asp	Kwas asparaginianowy
βCTF	Bezpośredni prekursor polipeptydów β -amyloidowych (ang. β C-terminal fragment of amyloid precursor protein)
CD	Spektroskopia dichroizmu kołowego
Co-IP	Koimmunoprecypitacja białek
DPPC	Dipalmitoilofosfatydylocholina
Cryo-EM	Kriomikroskopia elektronowa
Cys	Cysteina
DLS	Spektroskopia dynamicznego rozpraszania światła
DNA	Kwas deoksyrybonukleinowy
DSC	Skaningowa kalorymetria różnicowa
FRET	Transfer energii rezonansu fluorescencji
FLIM	Obrazowanie czasów życia fluorescencji
Glu	Kwas glutaminowy
Gly	Glicyna
Hsp70	Białka opiekuńcze z rodziny 70kDa
HPD	Motyw sekwencji domeny J białek pomocniczych zawierających domenę J zawierający histydynę, prolinę oraz kwas asparaginianowy
IR	Spektroskopia podczerwieni
ITC	Izotermiczna kalorymetria miareczkowa
JDP	Białka pomocnicze zawierające domenę J (ang. J-domain proteins)
Lys	Lizyna
MS	Spektrometria mas
NBD	Domena białka HSP70 wiążąca nukleotyd
NMR	Spektroskopia magnetycznego rezonansu jądrowego
PCA	Analiza głównych składowych
RAC	Kompleks białek opiekuńczych związanych z rybosomem (ang. Ribosome Associated Complex)
RMSD	Średnie odchylenie kwadratowe położenia
RNA	Kwas rybonukleinowy
SAXS	Spektroskopia niskokątowego rozpraszania promieniowania rentgenowskiego
SBD	Domena białka HSP70 wiążąca substrat
SOPC	1-stearoil-2-oleilofosfatydylocholina
SPR	Pomiar rezonansu plazmonów powierzchniowych
US	Technika umbrella sampling
Y2H	Drożdżowy system dwuhybrydowy
ZHD	Domena białka zuotyny (ang. Zuotin homology domain)



Rozdział 1

Wprowadzenie

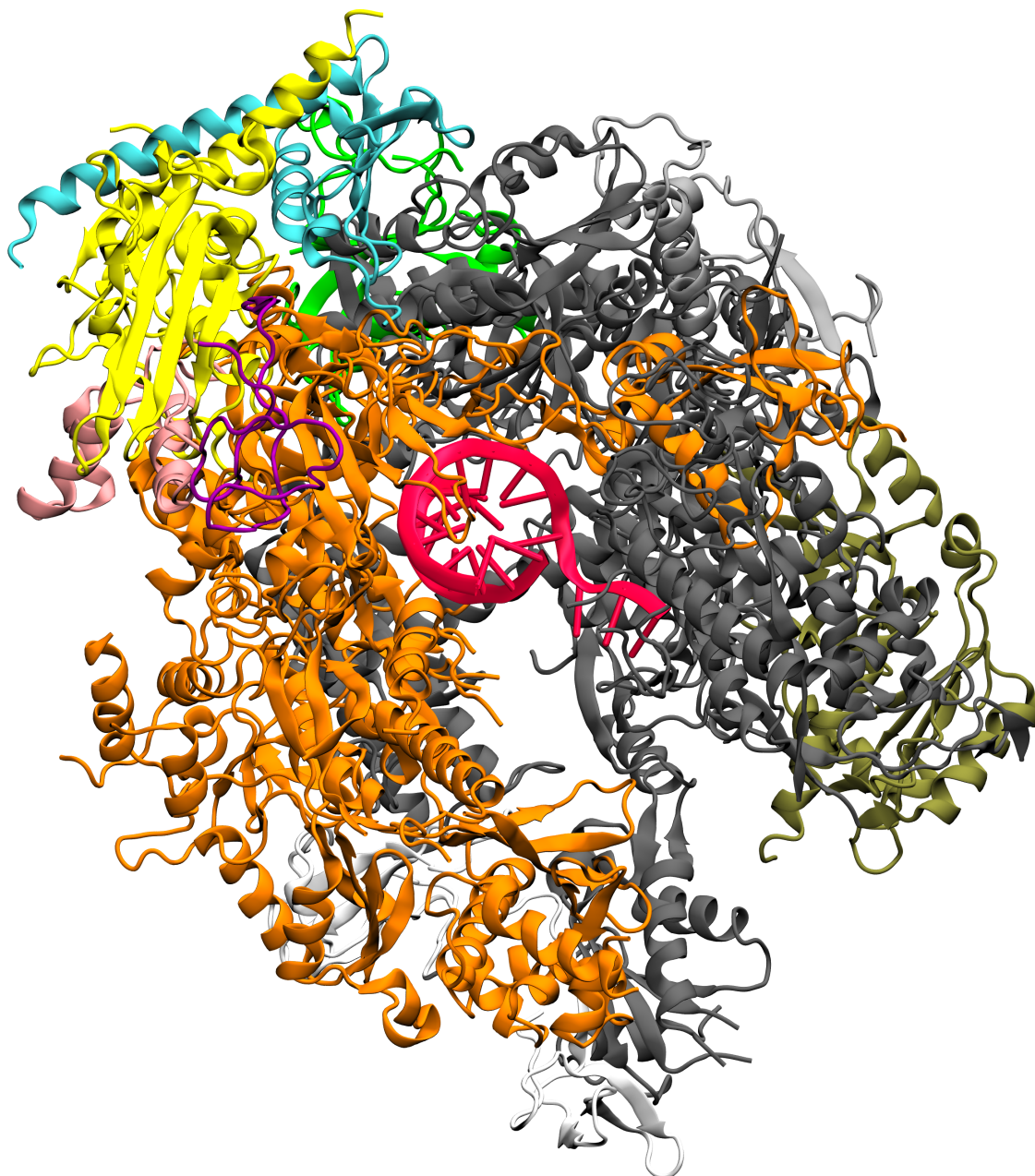
1.1 Oddziaływania białko-białko

Białka odgrywają kluczową rolę w praktycznie wszystkich procesach biologicznych, jednak rzadko kiedy działają w odosobnieniu. W zdecydowanej większości przypadków procesy te przeprowadzane są przez struktury makromolekularne, na które może składać się nawet kilkanaście różnych, specyficznie oddziałujących ze sobą białek, tak jak ma to miejsce w pokazanym na Rys. 1.1 kompleksie polimerazy II RNA. [1] W zależności od pełnionej funkcji, kompleksy białkowe mogą być złożone zarówno z białek tego samego rodzaju (homooligomery, np. szereg kanałów potasowych, takich jak Kv1.2 czy hERG, [2, 3] agregaty białkowe o zdefiniowanej funkcji, jak np. małe białka szoku termicznego [4]), jak i różnych rodzajów (heterooligomery, przede wszystkim bardzo wiele enzymów, np. syntaza ATP i karboksylaza rybulozo-1,5-bisfosforanu [5, 6]). Kompleksy białkowe mogą powstawać zarówno jako trwałe struktury, jak i tworzyć się tylko przejściowo, na potrzeby zrealizowania konkretnej funkcji. [7] By umożliwić pełnienie wybranej funkcji przez kompleksy białkowe, podjednostki tego kompleksu muszą znajdować się w ściśle zdefiniowanej orientacji względem siebie. Co więcej, białka powinny selektywnie oddziaływać z odpowiadającymi im partnerami wiązania, unikając tworzenia niespecyficznych kompleksów z innymi komponentami komórki. Taką specyficzną rozpoznania zapewniają wyspecjalizowane obszary wykształcone w strukturze białek, nazywanymi miejscami rozpoznania (ang. recognition site). Komplementarne miejsca rozpoznania białek pozwalają na utworzenie specyficznych powierzchni kontaktu między białkami, [8] a dzięki temu na selekcję partnerów wiązania. Podobnie jak centra aktywne, miejsca rozpoznania są zachowane ewolucyjnie, zaś ich zmiana często powiązana jest z nabywaniem nowej funkcji przez białko. [9]

1.2 Budowa miejsc rozpoznania białek

Łączna powierzchnia kontaktu pomiędzy dwoma komponentami kompleksu białkowego zwykle waha się pomiędzy 1200 a 2000 Å² (choć potrafi przekraczać nawet 3000 Å²), co odpowiada od 23 do 29 reszt aminokwasowych na każde z miejsc rozpoznania. [10] Reszty te zlokalizowane są zwykle w rejonach białka wyeksponowanych na zewnątrz, co ułatwia utworzenie sieci oddziaływań między nimi. [11] Wśród reszt występujących na powierzchni rozpoznania najczęściej spotyka się aminokwasy polarne: histydynę (His) i argininę (Arg), niepolarne: walinę, izoleucynę, leucynę, metioninę i cysteinę (Cys) oraz aromatyczne: fenyloalaninę (Phe), tyrozyne





RYSUNEK 1.1: Struktura polimerazy II RNA w kompleksie z RNA. Kolory odpowiadają kolejnym podjednostkom polimerazy. Łańcuch RNA zaznaczono kolorem malinowym.

(Tyr) i tryptofan (Trp). [12, 13] Dodatkowo, niektóre z aminokwasów są spotykane częściej w różnych rejonach miejsca rozpoznania. Z tej przyczyny powierzchnię miejsca rozpoznania dzieli się na obszar rdzeniowy (ang. core), złożony głównie z wcześniej wymienionych reszt, i krawędziowy (ang. rim), na którym znacznie częściej występują dodatkowo polarne aminokwasy: glicyna (Gly), asparagina (Asn) oraz kwas glutaminowy (Glu). [10] Aminokwasy te, poprzez różne typy oddziaływań (opisane poniżej), zarówno zapewniają siłę napędową dla asocjacji białek, jak i determinują ich wzajemną orientację w kompleksie.

1.3 Typy sił napędowych tworzenia kompleksu

W przypadku tworzenia kompleksów białkowych, dominującą rolę odgrywa rodzaj łańcuchów bocznych znajdujących się na powierzchni rozpoznania białka, w mniejszym zaś stopniu – rodzaj struktury drugorzędowej występującej na tej powierzchni. [14] Oddziaływania występujące między łańcuchami bocznymi aminokwasów można podzielić na kilka kategorii, w zależności zarówno od rodzaju oddziałujących reszt, jak i środowiska, w którym dochodzi do oddziaływań między białkami:

1.3.1 Oddziaływania polarne

Do oddziaływań powszechnie występujących na powierzchniach oddziaływań białkowych należą wiązania wodorowe, pary jonowe oraz mostki solne (pary jonowe tworzące wiązania wodorowe), występujące między hydrofilowymi grupami funkcyjnymi reszt aminokwasowych. [15] Kontakty tego typu obserwowane są zarówno bezpośrednio pomiędzy resztami białek, jak również mogą w nich uczestniczyć cząsteczki wody bądź jony. [16, 17] Z racji na dalekozasięgowy charakter oddziaływań elektrostatycznych, naładowane obszary na powierzchniach podjednostek kompleksów białkowych często odpowiadają za wzajemną preorientację ich miejsc rozpoznania jeszcze przed utworzeniem specyficznego kontaktu między resztami aminokwasowymi. [18] Rozpoznanie mające miejsce między naładowanymi domenami biomakromolekuł jest szczególnie widoczne na przykładzie kompleksów białko-DNA. [19] Dzięki wysokiej specyficzności tworzenia mostków solnych, często bywają one miejscami regulatorowymi na powierzchniach międzybiałkowych. [20]

Wśród naładowanych łańcuchów bocznych występujących na powierzchniach kontaktu białek najczęściej spotykane są reszty argininy. Tworzą one oddziaływania przede wszystkim z ujemnie naładowanymi Asp i Glu, jak również w mniejszym stopniu z hydrofilowymi łańcuchami bocznymi takimi jak np. Asn czy Gln. Arginina stanowi o tyle szczególny przypadek, iż dekamery poli(Arg) (ale nie poli(Lys)) są w stanie tworzyć struktury dimeryczne w środowisku wodnym pomimo silnego elektrostatycznego odpychania między nimi. [21] Stabilność takich dimerów wynika ze złożenia się kilku specyficznych cech grupy guanidynowej – poza posiadanym ładunkiem dodatnim jest ona zdolna tworzyć wiązania wodorowe z więcej niż jednym partnerem, a dodatkowo płaski kształt grup guanidynowych umożliwia tworzenie dodatkowych oddziaływań typu π -stacking (opisanych poniżej). [22] Pozwalają one na umiejscowienie kilku grup guanidynowych w jednym rejonie białka, zapewniając przeciwwagę dla odpychania elektrostatycznego między jednoimiennymi ładunkami.

1.3.2 π -stacking

Najpowszechniej spotykanymi łańcuchami bocznymi aminokwasów na powierzchniach oddziaływania białek są reszty aromatyczne (głównie Trp, Phe i Tyr), które tworzą kontakty przez oddziaływania typu π -stacking. [12] Bardzo często stanowią one tzw. kluczowe punkty kontaktu (ang. hot spots), które odpowiadają za istotny wkład do energii wiązania. [10] Oddziałujące ze sobą reszty ustawiają się

głównie w sposób równoległy względem siebie, co umożliwia efektywne nakładanie chmur elektronowych sprzężonych układów atomów o hybrydyzacji sp^2 . [23] W przeciwieństwie do dalej omawianego efektu hydrofobowego, wkład tworzonych kontaktów typu π -stacking do stabilizacji kompleksu dwóch molekuł w przeważającym stopniu zależy od siły ich oddziaływania ze sobą, [24] ta zaś w fazie gazowej niekiedy potrafi przeważać nawet korzyści płynące z utworzenia wiązań wodorowych. [25]

1.3.3 Efekt hydrofobowy

W przypadku niearomatycznych reszt o charakterze niepolarnym znajdujących się na interfejsie białkowym, ich główny wkład do stabilizacji kompleksu białkowego wynika z tzw. efektu hydrofobowego. Sugeruje się, że wynika on ze wzrostu entropii wysoce uporządkowanych cząsteczek wody znajdujących się w kontakcie z powierzchnią hydrofobową białka, które uwalniane są w wyniku asocjacji białek. [26] W pierwszym przybliżeniu skala tego efektu jest proporcjonalna do rozmiaru tworzonej powierzchni kontaktu między niepolarnymi łańcuchami bocznymi. [27] Zatem im lepsze dopasowanie powierzchni rozpoznania do siebie pod względem kształtu, tym większa siła napędowa wiązania skojarzona z efektem hydrofobowym. Zjawisko to wykorzystywane jest w inżynierii białek, gdzie w oparciu o nie możliwe jest zaprojektowanie powierzchni rozpoznania złożonych ze stosunkowo niewielkiej liczby aminokwasów, które będą wykazywały wysokie powinowactwo względem siebie, np. przy projektowaniu supramolekularnych struktur typu coiled-coil. [28]

1.3.4 Komplementarne powierzchnie kontaktu w błonie

Białka błonowe zawierają w swojej strukturze domeny złożone w zdecydowanej większości z aminokwasów o charakterze hydrofobowym, [29] które preferencyjne lokalizują się w hydrofobowym rdzeniu błony. [30] Rozpoznanie między tymi domenami zapewniają powierzchnie o komplementarnych kształtach, które umożliwiają utworzenie możliwie największej ilości korzystnych oddziaływania dyspersyjnych między sobą. [31] Jednym z przykładów takich powierzchni jest motyw GxxxG (gdzie G oznacza glicynę, zaś x - dowolny inny aminokwas), który obecny jest powszechnie w sekwencjach błonowych α -helis. [32–35] Jako że glicyna posiada wodór w miejscu łańcucha bocznego, motyw ten tworzy gładką powierzchnię umożliwiającą ścisły kontakt między białkami, a kontakt ten dodatkowo jest stabilizowany przez wiązania wodorowe tworzone pomiędzy atomami węgla $C\alpha$ glicyn a karbonylowymi atomami tlenu łańcucha głównego białek. [36, 37]

1.3.5 Niedopasowanie hydrofobowe

Poza specyficznym dopasowaniem powierzchni kontaktu, dodatkową siłą napędową tworzenia kompleksów białkowych w błonie jest tzw. niedopasowanie hydrofobowe (ang. hydrophobic mismatch). Ma ono miejsce w sytuacji, kiedy występuje różnica między grubością warstwy hydrofobowej błony a długością domeny białka osadzonej w tej błonie. [38] Skutkiem tego fragmenty hydrofobowe błony



ładź białka zostają wyeksponowane do wody, co jest niekorzystne termodynamicznie. Błona zatem wykazuje tendencję do ograniczania obszaru hydrofobowego wyeksponowanego na kontakt z wodą m.in. przez spychanie białek na siebie. [39–41] Można zatem traktować ten efekt jako analogiczny do efektu hydrofobowego w wodzie.

1.4 Techniki doświadczalne dla badania oddziaływań białko–białko

Bardzo duże zainteresowanie kontaktami międzybiałkowymi zaowocowało opracowaniem szeregu metod doświadczalnych pozwalających na badanie oddziaływań białko-białko na różnym poziomie szczegółowości. [42] Wśród nich możemy wyróżnić metody pozwalające na jakościową identyfikację występowania kontaktu między białkami, ilościowe oznaczenie ich wzajemnego powinowactwa, jak również strukturalne, dzięki którym możliwe jest wyznaczenie trójwymiarowej struktury tworzonych kompleksów białek.

1.4.1 Techniki jakościowe

Wśród metod jakościowego oznaczania oddziaływań białko-białko można wyróżnić m.in.:

- Drożdżowy system dwuhybrydowy (Y2H) [43]
- Koimmunoprecypitację białek (Co-IP) oraz precypitację kompleksów białkowych (ang. pull-down) [44, 45]
- Techniki western blot i far western blot [46, 47]
- Spektrometrię mas (MS) [48]
- Spektroskopia dynamicznego rozpraszania światła (DLS) [49]
- Spektroskopia niskokątowego rozpraszania promieniowania rentgenowskiego (SAXS) [50]
- Sיעiowanie białek (ang. cross-linking) przy pomocy odczynników spinających [51] bądź przy pomocy genetycznie wbudowanych aminokwasów z resztami fotoaktywnymi [52]

Pozwalają one na oznaczenie samego faktu tworzenia kompleksu (Y2H, Co-IP, pull down, oba warianty western blot) bądź na charakterystykę rozmiaru agregatów tworzonych przez białka (MS, DLS, SAXS). W przeciwieństwie do wcześniej wymienionych metod, sieciowanie białek pozwala na specyficzne wypróbkowanie reszt występujących w bezpośrednim kontakcie ze sobą. Choć niekiedy metody te prowadzą do oznaczania niespecyficznych kontaktów, stanowią nieocenioną pomoc w charakterystyce interfejsów białkowych. [53]

1.4.2 Techniki ilościowe

W przypadku technik ilościowych do najpopularniejszych technik należą:

- Skaningowa kalorymetria różnicowa (DSC) [54]
- Izotermiczna kalorymetria miareczkowa (ITC) [54, 55]
- Pomiar rezonansu plazmonów powierzchniowych (SPR) [56]
- Transfer energii rezonansu fluorescencji (FRET) w kombinacji z obrazowaniem czasów życia fluorescencji (FLIM) [57]

Dwie pierwsze techniki opierają się o pomiar efektów cieplnych związanych z zachodzącą asocjacją dwóch białek. W przypadku DSC wyniki uzyskuje się w oparciu o pomiar zmiany termogramu dla kompleksu względem izolowanych białek, zaś w przypadku ITC poprzez bezpośredni pomiar ciepła wydzielanego w wyniku dodawania kolejnych porcji jednego partnera wiązania do roztworu drugiego partnera. Obie techniki pozwalają zarówno na wyznaczenie energii swobodnej wiązania, jak również zmian entalpii i entropii związanych z procesem asocjacji. [54, 55] Szczególnie powszechnie wykorzystywana celem wyznaczania parametrów termodynamicznych jest technika ITC, która, choć bardzo użyteczna, ma pewne istotne wady: chociażby błędy w przygotowywaniu roztworów białek o dokładnym mianie stężenia skutkują dużymi błędami w wyznaczeniu wartości entalpii asocjacji białek, zaś te błędy dalej przenoszą się również na obliczaną wartość zmiany entropii. [58] Z tej też przyczyny nie jest to dobra metoda do wyznaczania powinowactwa dla bardzo stabilnych kompleksów, dla których stężenia roztworów białek wymagane do przeprowadzenia eksperymentu byłyby bardzo małe. W takich przypadkach DLC zwykle pozwala na wiarygodne oszacowanie parametrów termodynamicznych przez pomiar zmiany stabilności termicznej kompleksu względem monomerów.

W przypadku techniki SPR dokonuje się pomiaru zmiany współczynnika załamania światła na powierzchni badanej, gdzie jeden z partnerów wiązania jest immobilizowany na tej powierzchni, zaś drugi dostarczany jest jako roztwór w przepływie ciągłym. W ramach pomiaru bada się kolejno proces asocjacji partnera wiązania do białka związanego z powierzchnią, następnie występujący stan równowagi, i finalnie dysocjację partnera wprowadzając w przepływie wyłącznie roztwór buforowy. [56] Można więc z niego wyznaczyć zarówno stałą równowagi wiązania, jak również stałe szybkości asocjacji i dysocjacji.

Kombinacja technik FRET i FLIM również pozwala na pomiar stałej równowagi poprzez bezpośrednie wyznaczenie całkowitego stężenia dla każdego z partnerów wiązania z krzywych kalibracyjnych oraz stężenia kompleksu przez pomiar zaniku fluorescencji. [57] Metoda ta jest o tyle atrakcyjna, że umożliwia pomiary powinowactwa białek względem siebie *in vivo*.

1.4.3 Strukturalne

Wśród najpowszechniej stosowanych technik strukturalnych należy wyróżnić:

- Spektroskopia podczerwieni (IR) [59]
- Spektroskopia dichroizmu kołowego (CD) [60]
- Spektroskopia jądrowego rezonansu magnetycznego (NMR) [61]
- Rentgenografia strukturalna [62]
- Kriomikroskopia elektronowa (Cryo-EM) [63]

Dwie pierwsze techniki spektroskopowe stosowane są powszechnie przede wszystkim do określania struktury drugorzędowej białek, jednak same z siebie niosą jedynie ograniczoną informację strukturalną. [59, 60] W przeciwieństwie do nich, techniki NMR pozwalają na pomiar przesunięć chemicznych atomów o niezerowym spinie jądrowym oraz stałych sprzężenia między nimi. W oparciu o wielowymiarowe techniki NMR możliwe jest uzyskanie informacji zarówno o wiązaniach tworzonych między atomami w cząsteczce (np. HSQC [64]), jak również o występowaniu atomów blisko siebie w przestrzeni (np. NOESY [65]), co pozwala na ustalenia pełnej struktury białek bądź ich kompleksów. Rozdzielczość rejestrowanych sygnałów ogranicza jednak zastosowania tej metody w większości przypadków do białek o umiarkowanie małej masie (do ok. 50 kDa). [66]

Rentgenografia strukturalna również pozwala na scharakteryzowanie trójwymiarowej struktury białek w oparciu o pomiar dyfrakcji promieniowania wysokoenergetycznego zachodzącej na monokryształy białka. Technika ta pozwala w szczególnych przypadkach na śledzenie przebiegu reakcji chemicznych w białkach, indukowanych przy pomocy światła ultrafioletowego. [67, 68] Główną wadą tej metody jest jednak konieczność otrzymania kryształu białka dla wyznaczenia jego struktury. Dodatkowo, w niektórych przypadkach struktura białka bądź jego kompleksów może być zaburzona przez kontakty tworzone w sieci krystalicznej, które nie występują w warunkach natywnych w roztworze. [69, 70]

Stosunkowo nową techniką strukturalną jest Cryo-EM, w której wiązka elektronów rozpraszana jest na bardzo cienkiej warstwie badanego materiału - w przypadku białek najczęściej jest to ich zamrożony roztwór wodny (schłodzony na tyle szybko, by woda zachowała strukturę amorficzną). W wyniku pomiarów rozpraszanych elektronów zbiera się serie dwuwymiarowych rzutów gęstości elektronowej białek, z których możliwa jest rekonstrukcja pełnej struktury białka. Technika ta znacząco ułatwia wyznaczanie struktury białek błonowych, które można osadzić w odpowiednim detergencie. [71] Oprócz braku konieczności otrzymywania struktur krystalicznych, przewagą Cryo-EM nad krystalografią jest możliwość scharakteryzowania zestawu konformacji przyjmowanych przez białka w roztworze oraz brak tzw. problemu fazowego. [72] Technika zyskała w ostatnich latach tak duże uznanie, iż jej autorzy otrzymali za nią nagrodę Nobla w dziedzinie chemii w roku 2017. [73]

1.5 Metody obliczeniowe

Dzięki znaczącemu wzrostowi mocy obliczeniowej komputerów na przestrzeni ostatnich lat [74] metody modelowania komputerowego odgrywają coraz większą rolę w badaniu układów biomakromolekularnych. Obecnie metody komputerowe z powodzeniem pozwalają na przewidywanie struktury białek [75, 76], zmian konformacyjnych [77, 78] oraz, w przypadku małych białek, na badanie mechanizmów zwijania ich łańcuchów polipeptydowych, [79–81] dostarczając molekularnej interpretacji dla wyników otrzymywanych przez różnorodne techniki eksperymentalne. [82–84] Symulacje pełnoatomowej dynamiki molekularnej są również stosowane do śledzenia ścieżek tworzenia kompleksów dla małych domen białkowych bądź między sobą, bądź z DNA. [19, 85, 86] Śledzenie większych układów molekularnych do tej pory wymagało jednak rezygnacji z atomowej rozdzielczości na rzecz mniej szczegółowych metod opartych o uproszczone, gruboziarniste reprezentacje makromolekuł. [87–89]

1.6 Wybrane przykłady oddziaływań białko–białko

W niniejszej pracy pokazuje, że pełnoatomowy wariant symulacji dynamiki molekularnej może służyć zarówno do przewidywania struktur kompleksów białkowych, jak również dla scharakteryzowania sił napędowych leżących u podstaw ich asocjacji. By wykazać uniwersalność tych metod, do swoich symulacji wybrałem układy różniące się znacząco pod względem specyficzności oddziaływania białko–białko. Pierwszym z wybranych układów są białka opiekuńcze należące do rodziny 70–kDa (Hsp70), które oddziałują z białkami pomocniczymi należącymi do grupy białek z domeną J (JDP). [90] Układ ten jest wysoce interesujący z przynajmniej trzech powodów: jedną z nich jest powszechność występowania białek Hsp70 w organizmach żywych, zarówno prokariotów, jak i eukariotów. [91] Drugą przyczyną jest mnogość procesów, które są katalizowane przez białka Hsp70: do najbardziej znanych należy zwijanie białek, jednak w zależności od wiązanego partnera mogą one również brać udział w modyfikacjach białkowych czy translokacji białek przez błonę. [92] Trzecią przyczyną jest konserwacja reszt biorących aktywny udział w cyklach katalitycznych tych białek, zarówno w przypadku białka Hsp70, jak i współpracującego z nim JDP. [93] Udział tych reszt został już wielokrotnie potwierdzony na drodze eksperymentów biochemicznych. [94] Mimo to, aż do 2018 roku nie była znana struktura kompleksu podwójnego białka Hsp70 z białkiem JDP, [95] a struktura kompleksu potrójnego z substratem pozostaje nadal nierozwiązana. W ramach tej części badań rozważę dwa białka Hsp70: bakteryjne białko DnaK, katalizujące większość typowych procesów dla białek Hsp70, oraz drożdżowe białko Ssq1, które powstało w wyniku duplikacji mitochondrialnego homologa białka DnaK znanego u drożdży jako Ssc1 oraz o którym wiadomo, że wyspecjalizowało się wyłącznie do przenoszenia centrów żelazowo-siarkowych z miejsca ich syntezy w obręb białka rusztowania Isu1 na białka docelowe. Wybór pierwszego białka motywowany jest dostępnością struktury wyznaczonej przy pomocy rentgenografii strukturalnej zarówno dla samego białka DnaK, jak i dla kompleksu z domeną J białka JDP DnaJ. Wybór drożdżowego białka Ssq1 podyktowany był możliwością weryfikacji doświadczalnej hipotez stawianych w oparciu o metody komputerowe w ramach współpracy z zespołem prof. Marszałka z Międzyuczelnianego Wydziału Biotechnologii Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego.

Drugim rozważanym przeze mnie przypadkiem kontaktów białko–białko są oddziaływania w ramach 4-helikalnej domeny 4HB białka zuotyny. Białko to będące przedstawicielem białek JDP, współtworzy kompleks zasocjowany z rybosomem (ang. ribosome associated complex, RAC) spaja jednostki 60S i 40S rybosomu, [96] zaś sama domena 4HB oddziałuje z podjednostką 40S. Poza asocjacją z rybosomem, domena ta spełnia u części drożdży dodatkową rolę – aktywuje czynnik transkrypcyjny odpowiedzialny za ekspresję genów indukujących oporność wielolekową. [97] Funkcja ta wymaga, by domena 4HB uległa przynajmniej częściowemu rozwinięciu, prowadzącemu do ekspozycji C-terminalnego końca tej domeny.[98] W ramach tego systemu jest zatem możliwe prześledzenie zmian, jakie zachodziły na C-terminalnym fragmencie 4HB celem przystosowania tej domeny do pełnienia nowej funkcji oraz w jaki sposób zmiany te wpływały na sposób oddziaływania fragmentu C-końcowego z pozostałą częścią 4HB.

Ostatnim z wybranych układów jest enzym γ -sekretaza, który przeprowadza wewnątrz błonową proteolizę białek transbłonowych typu pierwszego. [99] W przeciwieństwie do specyficznego rozpoznania, występującego w przypadku białek

Hsp70 z białkami pomocniczymi JDP, dla γ -sekreazy znanych jest przynajmniej 90 różnych substratów, [100] wśród których dwoma najbardziej prominentnymi przykładami są białko Notch oraz prekursor polipeptydów β -amyloidowych. Zaburzone przetwarzanie białka Notch leży u podstaw rozwoju niektórych chorób nowotworowych, [101] z kolei zaburzenia w proteolizie prekursora polipeptydów β -amyloidowych uważane są za podłoże rozwoju choroby Alzheimera. [102] Choć enzym ten był intensywnie badany przez ostatnich 20 lat, jego struktura została rozwiązana dopiero w roku 2015 przy pomocy techniki kriomikroskopii elektronowej. [103] W literaturze można znaleźć zatem bardzo dużo danych biochemicznych, które można zastosować do weryfikacji wyników obliczeniowych. Ustalono, że działanie γ -sekreazy w znaczącym stopniu zależy od środowiska lipidowego, w którym się ono znajduje. Wiadomo, że zwiększenie zawartości cholesterolu w błonie wzmacnia intensywność proteolizy prowadzonej przez ten enzym. [104, 105] Wyniki te doprowadziły nawet do prób zastosowania terapii choroby Alzheimera opartej o statyny, które niestety zakończyły się fiaskiem. [106] Enzym ten stanowi zatem ciekawy przykład nie tylko ze względu na jego istotną rolę w chorobie Alzheimera, ale również ze względu na bezpośredni wpływ środowiska lipidowego na jego aktywność.

1.6.1 Budowa i funkcje białek rodziny HSP70

Nazwa białka szoku cieplnego (ang. Heat Shock Protein) została nadana grupie białek, których ekspresja wzrasta, gdy komórki poddawane są warunkom stresowym, jak np. podwyższona temperatura czy ciśnienie osmotyczne. [107] Hsp70 wykorzystują energię uwalnianą w wyniku hydrolizy ATP do cyklicznego wiązania i uwalniania krótkich hydrofobowych fragmentów łańcuchów innych białek, ułatwiając w ten sposób przeorganizowanie ich struktury drugorzędowej. [108] Dzięki tej zdolności białka Hsp70 biorą udział w wielu procesach komórkowych: związane z rybosomem pośredniczą w zwijaniu nowo syntezowanych łańcuchów polipeptydowych, wspomagają odzyskanie właściwej struktury przez częściowo zdenaturowane białka, mediują oddziaływania między niektórymi białkami czy napędzają transport łańcuchów polipeptydowych przez błony biologiczne. [92, 109] Katalizę prowadzoną przez białka Hsp70 ukierunkowuje każdorazowo jedno z wielu białek pomocniczych typu-J, wyspecjalizowanych dla każdego z wymienionych procesów. [110]

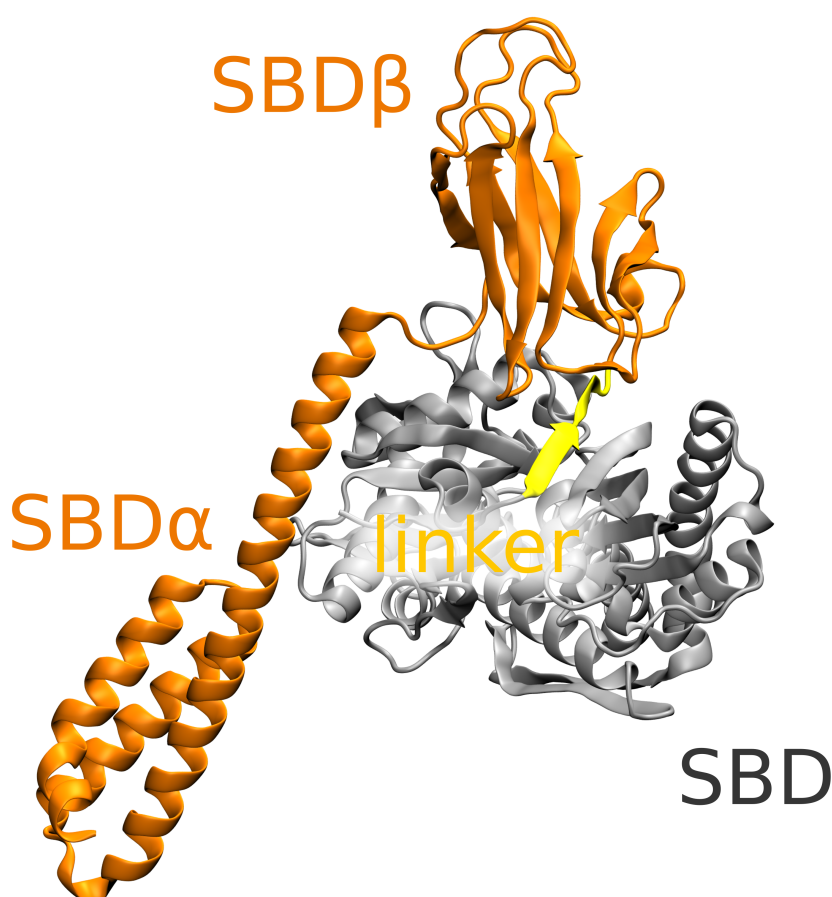
Białka HSP70 zbudowane są z dwóch domen (Rys. 1.2): domeny wiążącej substrat (ang. Substrate Binding Domain, SBD) oraz domeny wiążącej nukleotyd (ang. Nucleotide Binding Domain, NBD). SBD składa się z dwóch fragmentów – SBD- α , złożonego głównie z alfa-helis, oraz SBD- β o strukturze β -beczki. Gdy substrat zostaje związany przez białko HSP70, fragment SBD- α zamyka się nad substratem związanym do SBD- β , zmieniając konformację domeny na zamkniętą oraz uniemożliwiając oddysocjowanie polipeptydowego fragmentu. [111] Podobnie otwarcie domeny wiąże się z oddysocjowaniem fragmentu helikalnego SBD- α , uwalniając związany fragment peptydowy oraz pozwalając na kontynuację cyklu katalitycznego.

Domena NBD swoją budową przypomina heksokinazy, w których cztery subdomeny białka (IA, IB, IIA, IIB) układają się naokoło wnęki wiążącej nukleotyd. [112] Nukleotyd lokuje się we wnęcie pomiędzy dwoma β -spinkami, na których znajduje się treonina biorąca udział w hydrolizie. Domeny te są ściślej zasocjowane w przypadku związania ATP. [113] W stanach ze związanym ADP bądź bez związanego

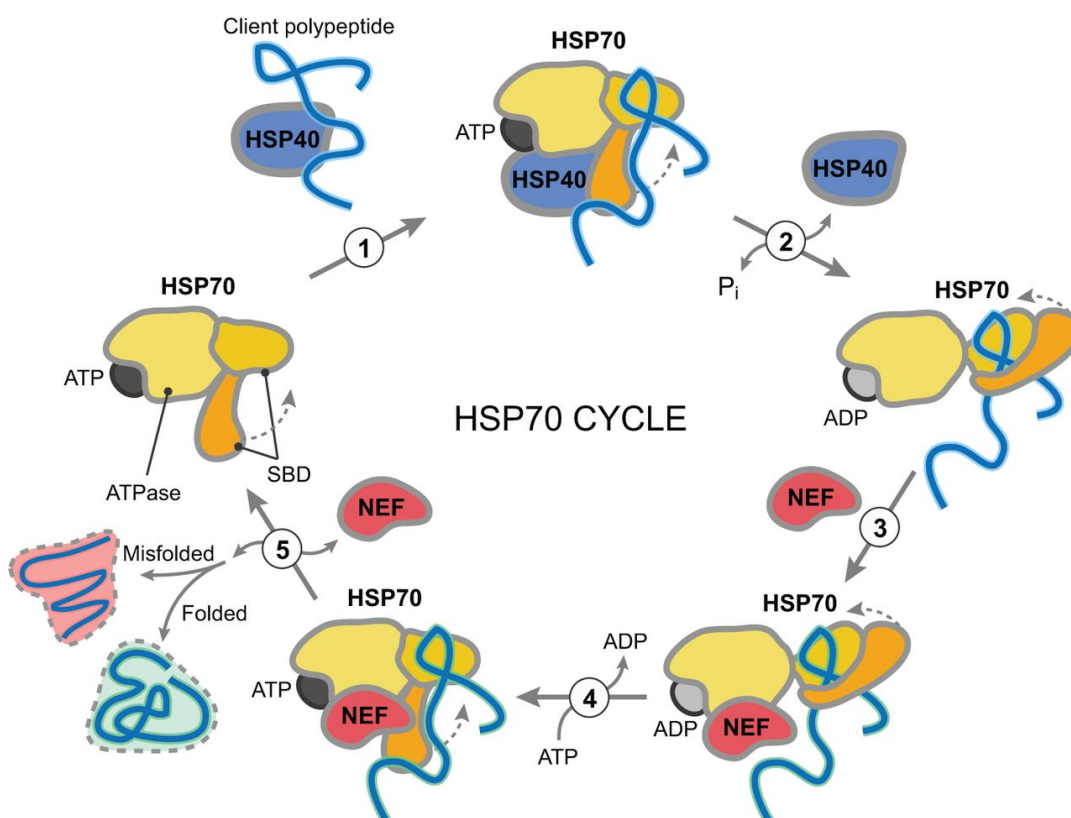
nukleotydu, płat tworzony przez subdomeny IA i IB odchyła się od drugiego płata, składającego się z subdomen IIA i IIB. [114]

Konformacja domeny SBD zależna jest od rodzaju nukleotydu związanego przez NBD. Gdy w kieszeni wiążącej nukleotydu NBD występuje ATP, SBD- α występuje głównie w formie zasocjowanej do NBD, zachowując konformację otwartą domeny SBD. [113] Z kolei po hydrolizie ATP przez NBD enzym występuje w konformacji, w której NBD i SBD są rozdysocjowane, a SBD przyjmuje konformację zamkniętą (z SBD- α przykrywającym kieszeń SBD- β wiążącą substrat). [115]

Cykl katalityczny białka Hsp70, oparty o allosteryczną komunikację między domenami SBD i NBD, przedstawiony został na Rys. 1.3. [116] Jako pierwsze następuje związanie białka typu-J, które dostarcza substrat do SBD. Następnie ATP ulega hydrolizie, a białko J oddysocjowuje od kompleksu z Hsp70. Na tym etapie również następuje zamknięcie się domeny SBD. Kolejnym krokiem jest wymiana nukleotydu z ADP na ATP. W przypadku niektórych białek Hsp70 ADP ulega samoistnemu oddysocjowaniu, [117] jednak w większości przypadków wymiana ta jest stymulowana przez dodatkowe białko pomocnicze – czynnik wymiany nukleotydu. [118] Kończącym etapem jest otwarcie SBD pod wpływem związania ATP przez NBD, co



RYSUNEK 1.2: Budowa białek rodziny Hsp70. Kolorem szarym zaznaczono domenę wiążącą nukleotydu, pomarańczowym – domenę wiążącą substrat, zaś żółtym β -nić łączącą domeny (ang. linker).



RYSUNEK 1.3: Cykl katalityczny białka Hsp70. Schemat wzięty z pracy [116]

umożliwia rozpoczęcie kolejnego cyklu wiązania substratu.

Szlak allosteryczny leżący u podstaw komunikacji między domenami SBD i NBD nie jest znany, jednak do tej pory udało się zidentyfikować przynajmniej część reszt biorących w nim udział. Wśród nich, jedną z lepiej poznanych reszt jest arginina, (Arg167 w bakteryjnym DnaK bądź Arg207 w drożdżowym Ssq1), która znajduje się na NBD i oddziałuje bezpośrednio z SBD. [94] Mutant DnaK Arg167Ala hydrolizuje ATP z szybkością porównywalną do białka dzikiego stymulowanego zarówno przez białko typu-J, jak i substrat. [113] Dodatkowo, również w przypadku podwójnego mutantu Arg167Cys Gly480Cys DnaK obserwowano podwyższoną intensywność hydrolizy ATP względem białka dzikiego oraz białko to nie podlegało stymulacji przez obecność białka JDP DnaJ. [113] Co jednak zaskakujące, utlenienie tych reszt do wiązania disiarczkowego, choć zapobiega pełnemu rozdysocjowaniu domen NBD i SBD, prowadziło do poziomu stymulacji hydrolizy podobnej, jak w przypadku mutantu Arg167Ala. [113] Wynik ten sugeruje, że pełne rozdysocjowanie domen białka DnaK nie jest konieczne dla hydrolizy ATP, a wymagana jest raczej zmiana interfejsu oddziaływania domen SBD i NBD.

Dotychczasowe wyniki sugerują, że stymulacja hydrolizy białek Hsp70 poprzez białka typu-J opiera się o oddziaływanie z wcześniej wspomnianą arginina. W przypadku białek typu-J kluczowy dla ich aktywności stymulacyjnej jest motyw HPD, występujący między helisami domeny J. Pokazano, iż mutacja którejkolwiek z reszt motywu histydyna, prolina, kwas asparaginowy (HPD) drastycznie obniża aktywność stymulacyjną białka DnaJ. [119] Co więcej, pokazano, że białko DnaJ z mutacją Asp35Asn w obrębie motywu HPD stymuluje białko DnaK, jeśli wprowadzić

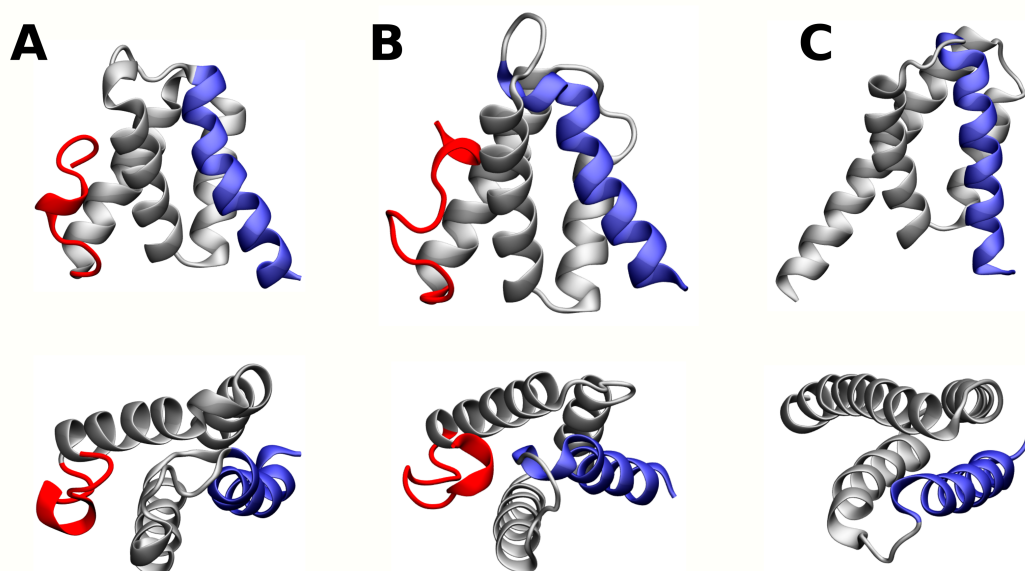
na nim mutację Arg167His. [94] Podobnie, mutacje motywu HPD okazały się letalne dla *Schizosaccharomyces pombe*. [120] Ku zaskoczeniu jednak, w przypadku wysoce wyspecjalizowanej pary białka J Jac1 i białka Hsp70 Ssq1 u *Saccharomyces cerevisiae*, pojedyncze mutacje motywu HPD osłabiały wzrost drożdży, ale nie były letalne. [120] W tym przypadku obszarem dodatkowo odpowiedzialnym za stymulowanie aktywności białka Hsp70 okazała się helisa 2 domeny J, zawierająca cztery dodatkowo naładowane reszty Lys i Arg. Otrzymana rentgenograficznie w 2018 roku struktura białka DnaK w kompleksie z DnaJ pokazała, że reszty Arg167 DnaK i Asp35 DnaJ oddziałują bezpośrednio ze sobą w kryształach. [95] Pozostaje jednak niejasnym, w jaki sposób to oddziaływanie wpływa na postęp cyklu katalitycznego.

1.6.2 Budowa i funkcje białka zuotyny

Białko zuotyna należy do grupy białek z domeną J, które poprzez współpracę z białkami HSP70 ułatwiają prawidłowe zwijanie nowo zsyntetyzowanych łańcuchów polipeptydowych. [121] Wraz z białkiem z rodziny HSP70 niezdolnym do hydrolizy ATP współtworzy ono kompleks zasocjowany z rybosomem (ang. ribosome associated complex, RAC), który funkcjonuje w zwijaniu nowo powstałych łańcuchów polipeptydowych. [122] RAC wiąże się z zarówno podjednostkami 60S jak i 40S rybosomu, [123] pozycjonując homologiczną domenę zuotyny (ang. zuotin homology domain, ZHD) w pobliżu wyjścia kanału podjednostki 60S rybosomu. [124] Dzięki temu RAC, poprzez oddziaływanie ze swoim partnerem HSP70 (np. Ssb1 u drożdży), lokalizuje go w bezpośrednim sąsiedztwie łańcucha polipeptydowego uwalnianego z rybosomu, ułatwiając jego zwijanie. [121]

Drożdzowe białko zuotyna składa się kolejno z domeny N-końcowej, domeny J, domeny ZHD, naładowanego regionu oraz domeny 4HB. [125] W przypadku roślin, zwierząt i pierwotniaków, na C-końcu homologów białka zuotyny znajdują się również domeny SANT, [126, 127] które odpowiadają m.in. za oddziaływanie z chromatyną. [128] Pomimo różnic strukturalnych wiązanie zuotyny do rybosomu występuje zarówno u drożdży, jak i u pozostałych gatunków, które na C-końcu posiadają domeny SANT. Co więcej, wariant ludzki zuotyny z powodzeniem zastępował natywne białko u drożdży. [129] Struktura kompleksu rybosomu z RAC pokazuje, że RAC łączy się z jednostką 40S rybosomu poprzez 4HB, [123] zaś badania biochemiczne wskazały na istotną rolę dodatnio naładowanej helisy 1 4HB w tym rozpoznaniu.

Poza oddziaływaniem z rybosomem, zuotyna pełni również dodatkowe funkcje, m.in. moduluje oddziaływanie białek z chromatyną. [130] W przypadku drożdży, zuotyna aktywuje czynnik transkrypcyjny odpowiedzialny za ekspresję genów indukujących oporność wielolekową. [97] Za tę aktywację u *Saccharomycetaceae* odpowiada 4HB, a dokładnie 13 reszt znajdujących się na jej C-terminalnym końcu. [98] U drożdży fragment ten tworzy nieustrukturyzowany łańcuch wychodzący z helisy 4, który asocjuje do szczeliny między helisami 1,2 a 3 i 4 4HB (Rys. 1.4). [98, 125] Dla aktywacji czynnika transkrypcyjnego nie wystarczy zatem, by 4HB oddysocjowała od rybosomu. Dodatkowo musi również ulec przynajmniej częściowemu rozwinięciu, co umożliwi ekspozycję jej C-terminalnego fragmentu do partnera wiązania. Oszacowana stabilność tej domeny u *Saccharomyces cerevisiae* jest rzeczywiście dość niska – od około -1 do -2 kcal/mol. Można podejrzewać, że zachodzące zmiany ewolucyjne indukowały zarówno przysposobienie 4HB do oddziaływania z czynnikami aktywującymi ekspresję, jak również przyczyniły się



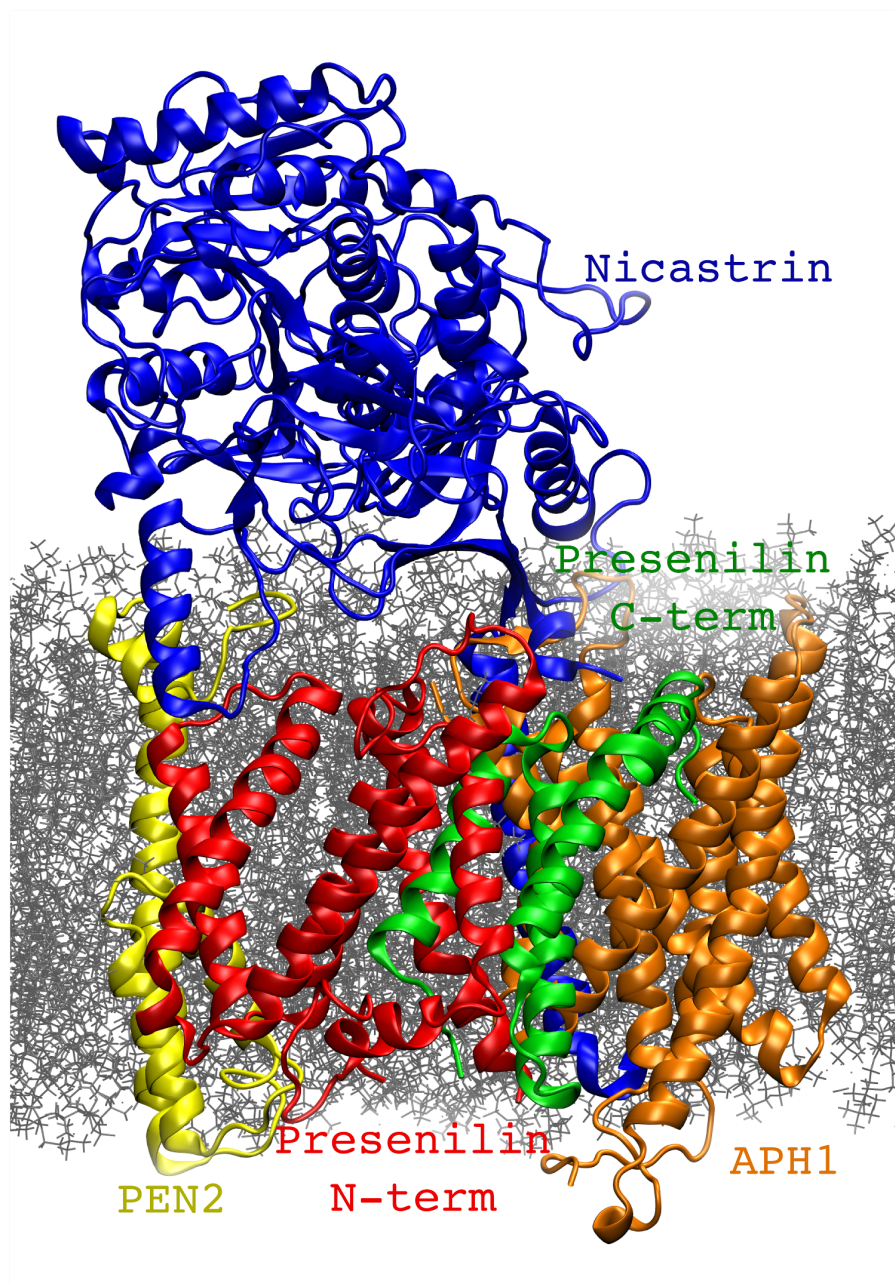
RYSUNEK 1.4: Struktura 4HB organizmów drożdżowych (A *Saccharomyces cerevisiae* i B *Chaetomium thermophilum*) oraz C *Homo sapiens*. Kolorem niebieskim zaznaczono helisę 1, odpowiadającą za wiązanie do rybosomu, zaś kolorem czerwonym 13 C-terminalnych reszt wystających poza helisę 4.

do jej częściowej destabilizacji. Wykazywanie tendencji C-terminalnego końca do zawijania się do szczeliny między helisami 4HB nie jest obserwowane w przypadku białka ludzkiego ani nie jest przewidywane w przypadku innych organizmów zwierzęcych.[131] Nasuwa to pytanie, czy C-terminalny koniec 4HB wpływa istotnie na stabilność tej domeny w przypadku jej drożdżowych wariantów.

1.6.3 Budowa i funkcje γ -sekreazy

γ -sekreaza jest enzymem błonowym biorącym udział w procesach kontrolowanej proteolizy wewnątrz błonowej. Na ten moment znanych jest przynajmniej 90 różnych substratów tego enzymu, [100, 132] a do najbardziej rozpoznawanych należą białka β CTF oraz Notch. β CTF jest bezpośrednim prekursorem polipeptydów β -amyloidowych, których nadmierna sekrecja prowadzi do rozwoju choroby Alzheimera, dotykającej około 10 % społeczeństwa w wieku powyżej 65 lat. [133] Zaburzenia w sygnalizacji Notch z kolei prowadzą do rozwoju niektórych chorób nowotworowych. [134]

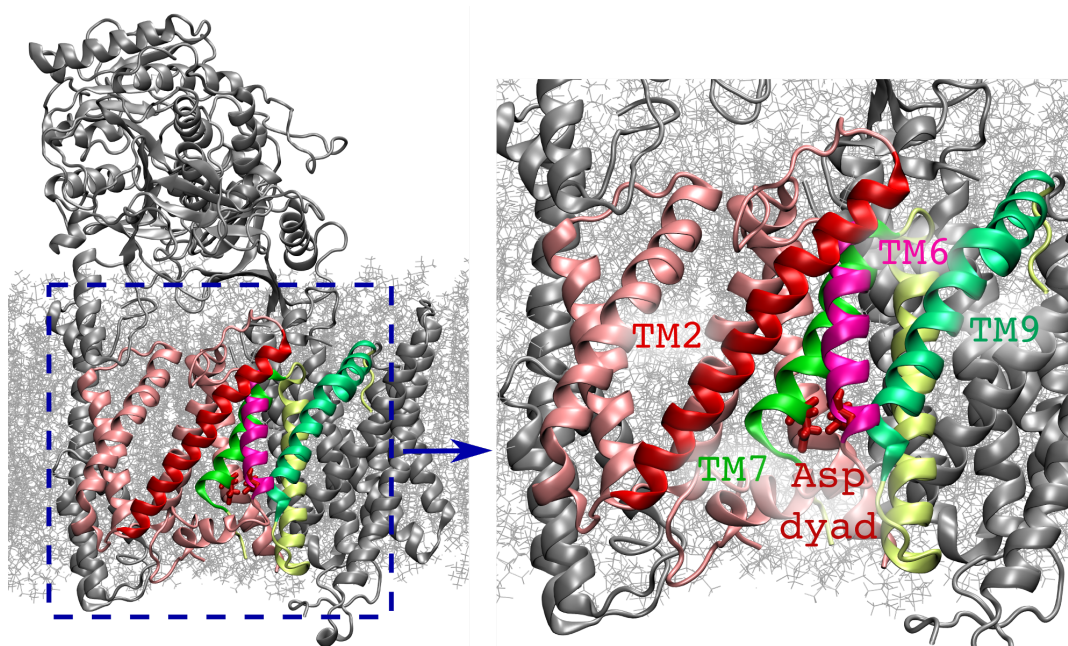
γ -sekreaza jest heterotetramerem złożonym z 4 białek: nikastryny, APH-1, PEN-2 oraz preseniliny (Rys. 1.5). Presenilina, będąca katalityczną podjednostką γ -sekreazy, jest proteazą aspartylową, złożoną z 9 transbłonowych α -helis. Katalityczne reszty asparaginianowe (Asp257 i Asp385) znajdują się w obrębie helis 6 i 7 (Rys. 1.6). γ -sekreaza jest aktywowana w wyniku autoproteolizy dzielącej presenilinę na dwa fragmenty: N-końcowy, składający się z pierwszych 6 helis, oraz C-końcowy, złożony z pozostałych 3 helis. Struktura krystaliczna przodka preseniliny (białka PSH) pokazała, że w stanie wolnym reszty aspartylowe znajdują się w znaczącej odległości od siebie (6,7 Å). [135] Sugeruje to, że w presenilinie zmiana konformacyjna umożliwiająca proteolizę powinna zachodzić dopiero pod wpływem



RYSUNEK 1.5: Struktura kompleksu γ -sekreazy umieszczonego w błonie lipidowej.

wiązania substratu. Dodatkowo, zaobserwowano, że w γ -sekreazie występują dwa miejsca wiązania substratu – wstępne oraz katalitycznie aktywne. [136] Miejsca te ułożone są prawdopodobnie w nieznacznej odległości od siebie, [136] co sugeruje, że znajdują się one w obrębie preseniliny. Dostęp do centrum aktywnego enzymu kontrolowany jest przez szczelinę między helisami, zwaną „bramą boczną” (ang. lateral gate). [137]

Dotychczas sugerowano, iż za dostęp do centrum aktywnego może odpowiadać helisa 9 preseniliny, która wydawała się także współtworzyć miejsce wstępnego wiązania substratu. [138] Wyniki otrzymane metodą kriomikroskopii elektronowej wskazują jednak na dużą dynamikę konformacyjną helisy 2, która najprawdopodobniej wykazuje tendencję do oddysocjowywania od reszty kompleksu, otwierając dostęp

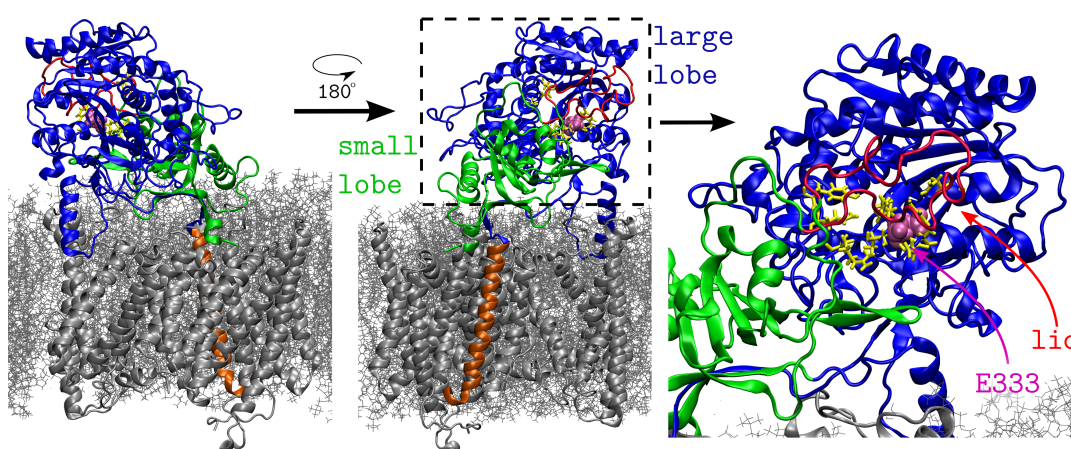


RYSUNEK 1.6: Budowa preseniliny - katalitycznej jednostki γ -sekretazy.

do centrum katalitycznego. [139] Otrzymana struktura wskazuje również na istotną rolę helisy 9 dla formowania kompleksu między fragmentem C-terminalnym preseniliny a podjednostką APH-1. [103] Każde to sądzić, iż helisa 9 może raczej odpowiadać za odpowiednie usytuowanie obu tych podjednostek względem siebie niż za dostęp substratu do centrum aktywnego enzymu. Zgodnie z ostatnio otrzymanymi strukturami kompleksów γ -sekretazy z substratami Notch i β CTF, [140, 141] dostęp do centrum katalitycznego od strony helisy 6 i 9 ograniczony byłby przez bardzo krótki łącznik między helisami 5 i 6 preseniliny, co zdaje się dodatkowo przemawiać za wiązaniem substratów przez enzym od strony helisy 2 i 6. Dostęp do obu proponowanych obszarów, mających kontrolować dostęp do centrum aktywnego enzymu, wymaga wstępnego zadokowania substratu w bezpośrednim sąsiedztwie helisy 6.

O ile przyjmuje się zwykle, że podjednostki APH-1 i PEN-2 biorą udział w stabilizacji i dojrzewaniu struktury preseniliny, [142] o tyle rola nikastryny pozostaje niejasna. Nikastryna jest białkiem transbłonowym typu pierwszego, z pojedynczą helisą transbłonową oraz z zaskakująco dużym, globularnym fragmentem zewnątrzkomórkowym, który stanowi około 45% całkowitej masy γ -sekretazy. Fragment zewnątrzkomórkowy nikastryny składa się z dwóch „płatów” (ang. lobe): małego, ulokowanego na N-końcu białka oraz obejmującego 5 helisy oraz 10 β -nici, oraz dużego, znajdującego się pomiędzy małym płatem a helisą transbłonową i obejmującego 12 helisy oraz 14 β -nici. [143] W skład małego płata wchodzi pętla, która tworzy wieko osłaniające potencjalną kieszeń wiążącą fragment N-terminalny substratu (Rys. 1.7). W oparciu o strukturę homologów nikastryny zasugerowano, iż kieszeń ta miałaby być odsłaniana w wyniku obrotu płatów względem hydrofobowego trzpienia, znajdującego się w centralnym obszarze styku płatów. [143] Rola nikastryny jako elementu rozpoznającego substrat budzi jednak wciąż kontrowersje. Wyniki koimmunoprecypitacji wskazują bowiem, że izolowany fragment zewnątrzkomórkowy nikastryny wykazuje tendencję do wiązania peptydowych substratów

γ -sekreazy. [144] Szczególnie istotna dla tego wiązania wydaje się być wysoce zakonserwowana reszta Glu333, która miałaby pomagać w rozpoznaniu dodatnio naładowanego N-terminalnego końca substratów. Pokazano, że mutacje Glu333 do alaniny inhibowały aktywność γ -sekreazy względem substratów przy zachowaniu aktywności katalitycznej samej preseniliny. Odpowiednio, usunięcie ładunku dodatniego z N-końca białka β CTF przez formylowanie również hamowało proces proteolizy. [144] Ostatnie wyniki biochemiczne wskazują jednak, że aktywność proteolityczna γ -sekreazy względem C-terminalnego fragmentu białka Notch jest niezależna od oddziaływania z jego N-końcem. [145] Co więcej, modyfikacje tego substratu przez wydłużanie bądź skracanie jego N-końca pozwalały na odpowiednio zmniejszenie bądź zwiększenie aktywności katalitycznej γ -sekreazy, zaś usunięcie wiązań disiarczkowych z zewnątrzkomórkowej domeny nikastryny w aktywnym kompleksie ułatwiało proteolizę wydłużonych fragmentów białka Notch. [145]



RYSUNEK 1.7: Struktura nikastryny. Na prawym panelu kolorem żółtym zaznaczono hydrofilowe reszty potencjalnej kieszeni wiążącej substrat, kolorem czerwonym wieko osłaniające tę kieszeń, zaś kolorem fioletowym resztę Asp333, sugerowaną jako kluczową dla rozpoznania substratów.

Pozwoliło to na zaproponowanie alternatywnego mechanizmu, zgodnie z którym nikastryna stanowi element wykluczający przestrzennie białka o dużych domenach zewnątrzkomórkowych, co zapobiegałoby ich proteolizie przez presenilinę. [145] Zgodnie z tą koncepcją, głównym motorem napędowym dla wiązania substratu jest korzystne oddziaływanie transbłonowych helis preseniliny z substratem peptydowym.

Aktywność γ -sekreazy silnie zależy od środowiska lipidowego, w którym się znajduje. Do tej pory pokazano, że aktywność tego enzymu rośnie wraz ze wzrostem zawartości cholesterolu w błonie. [104, 105] Usunięcie cholesterolu z błony komórkowej całkowicie hamowało aktywność enzymu. [104] Na tej podstawie zasugerowano, iż γ -sekreaza lokalizuje się w tzw. tratwach lipidowych - ciekłych uporządkowanych mikrodomenach błony, wzbogaconych w cholesterol i sfingolipidy. Co więcej, w przypadku białek β CTF i Notch pokazano również, że wykazują one specyficzne oddziaływania z cholesterolem, [146, 147] mogące preferencyjnie lokować je w takich tratwach. Jednym z proponowanych mechanizmów stymulacji aktywności γ -sekreazy przez cholesterol miałyby być ściślejsze związanie N-terminalnych

fragmentów substratów do błony, co zmniejszałoby ich ekspozycję na oddziaływania odpychające z nikastryną. [145] Inne prace sugerują bezpośrednią rolę kompleksów cholesterol-substrat w efektywnym wiązaniu do enzymu. [146] Mechanizmy te jednak nie tłumaczą braku aktywności enzymu pod nieobecność cholesterolu.

Rozdział 2

Cele oraz zakres pracy

W ramach niniejszej pracy pokazuję, w jaki sposób metody symulacji komputerowych mogą pomóc w identyfikacji mechanizmów rozpoznania i wiązania pomiędzy białkami. By uwypuklić uniwersalność tych metod, stosowane techniki zostały użyte zarówno do badania układów homogenicznych (białka w roztworze wodnym), jak i heterogenicznych (białka osadzone w dwuwarstwie lipidowej).

W przypadku białek rodziny HSP70 planuję ustalić mechanizm rozpoznania pomiędzy drożdżowym białkiem Ssq1 a białkiem pomocniczym Jac1. W pierwszej części badań, opierając się o wcześniejsze analizy ewolucyjne [120] planuję zweryfikować, czy dodatnio naładowanemu obszarowi domeny J białka Jac1 odpowiada komplementarny, ujemnie naładowany obszar na powierzchni białka Ssq1. Następnie sprawdzę, czy te naładowane regiony mogą służyć do utworzenia produktywnego kompleksu. Niezależnie, w oparciu o metody przeszukania przestrzeni konfiguracyjnej, scharakteryzuję występowanie różnych modów wiązania kompleksu Jac1-Ssq1. Na zakończenie tej części planuję oszacować wzajemne powinowactwo białek, a swoje przewidywania zweryfikuję poprzez scharakteryzowanie zmiany powinowactwa białka Jac1 pod wpływem podwójnej mutacji Arg37Ala Arg41Ala. Doświadczalnie pokazano, że taki mutant Jac1 nie wiąże się do Ssq1.

W kolejnej części badań chcę ustalić w jaki sposób związanie Jac1 stymuluje postęp cyklu katalitycznego Ssq1. W tym celu chcę ustalić w jaki sposób związanie białka Jac1 wpływa na zmianę interfejsu oddziaływań między domenami NBD i SBD białka Ssq1 ze szczególnym uwzględnieniem roli motywu HPD w indukowaniu tej zmiany. By sprawdzić czy wyniki uzyskane dla Ssq1 mają zastosowanie dla innych przedstawicieli rodziny Hsp70, badania te powtórzę również dla kompleksu białka DnaK z domeną J białka DnaJ, którego struktura krystaliczna została opublikowana trakcie realizacji tej pracy doktorskiej. [95] Na zwieńczenie tej części badań, planuję zweryfikować hipotezy dotyczące roli białek JDP w postępie cyklu katalitycznego białek HSP70 w oparciu o eksperymentalne dane mutagenyzy białek HSP70. [113, 115, 148]

W ostatniej części poświęconej białkom HSP spróbuję wyjaśnić, w jaki sposób związanie domeny J przez HSP70 indukuje zmianę interfejsu NBD-SBD. Najpierw na przykładzie kompleksu DnaK-DnaJ zamierzam ustalić, jaką rolę odgrywa asparaginian należący do motywu HPD w rozpoznaniu pomiędzy tymi białkami. W kolejnym kroku postaram się scharakteryzować jakie elementy domeny DnaJ są konieczne dla zmiany interfejsu NBD-SBD. Sumarycznie, otrzymane wyniki pozwolą na scharakteryzowanie roli białek pomocniczych w cyklu katalitycznym białek HSP70.

W ramach badań dotyczących domeny 4HB zuotylny planuję sprawdzić, jakie zmiany na C-terminalnym fragmencie 4HB umożliwiły jej do pełnienie nowej funkcji. W tym celu na początku porównam różnice strukturalne domen 4HB pomiędzy wariantem ludzkim a drożdżowym. Następnie, w oparciu o modele homologiczne ancestralnych wariantów 4HB prześlę, w jaki sposób zmieniały się oddziaływania między helisami domeny a jej C-terminalnym fragmentem. W kolejnym kroku sprawdzę, jak zmieniała się stabilność 4HB w trakcie ewolucji oraz w jakim stopniu zmieniał się udział C-terminalnego końca w stabilizacji domeny. Finalnie scharakteryzuję dynamikę konformacyjną 4HB *Saccharomyces cerevisiae*, u których stymuluje ona oporność wielolekową.

W przypadku badań dotyczących γ -sekreazy planuję sprawdzić, w jaki sposób środowisko lipidowe reguluje aktywność tego enzymu wobec białka β CTF jako przykładowego substratu. W tym celu na początku zbadam, czy obecność cholesterolu wpływa na powinowactwo substratu do wstępnego miejsca wiązania enzymu osadzonego w błonie złożonej z dipalmitoilofosfatydylocholiny (DPPC). Następnie sprawdzę, czy obserwowana różnica powinowactwa może wynikać z niedopasowania hydrofobowego (ang. hydrophobic mismatch) [38] pomiędzy błoną a substratem bądź enzymem. Uzyskane wyniki pozwolą na zaproponowanie roli cholesterolu w rozpoznaniu substratów przez γ -sekreazę.

Rozdział 3

Metody obliczeniowe

3.1 Elementy mechaniki statystycznej

Mechanikę statystyczną stosuje się do opisu układu składającego się z wielu ciał (cząstek). Podstawowym postulatem mechaniki statystycznej jest takie samo prawdopodobieństwo występowania zdegenerowanych stanów o tej samej energii (E), objętości (V) oraz liczbie cząstek (N). Powyższy postulat pozwala na zdefiniowanie entropii jako miary liczby stanów mikroskopowych układu (mikrostanów) odpowiadających makrostanowi określonego przez wartości N , V i E :

$$S(N, V, E) = k_B \ln W(N, V, E) \quad (3.1)$$

gdzie przez $S(N, V, E)$ oznaczamy entropię danego makrostanu, k_B oznacza stałą Boltzmanna, zaś $\Omega(N, V, E)$ liczbę mikrostanów realizujących makrostan o zadanych parametrach N , V i E . Entropia jest funkcją charakterystyczną układu o stałej wartości energii, objętości i liczbie cząstek, nazywanego zespołem mikrokanonicznym. Odpowiada on sytuacji, w której układ jest izolowany od otoczenia (nie wymienia z nim ani energii, ani liczby cząstek).

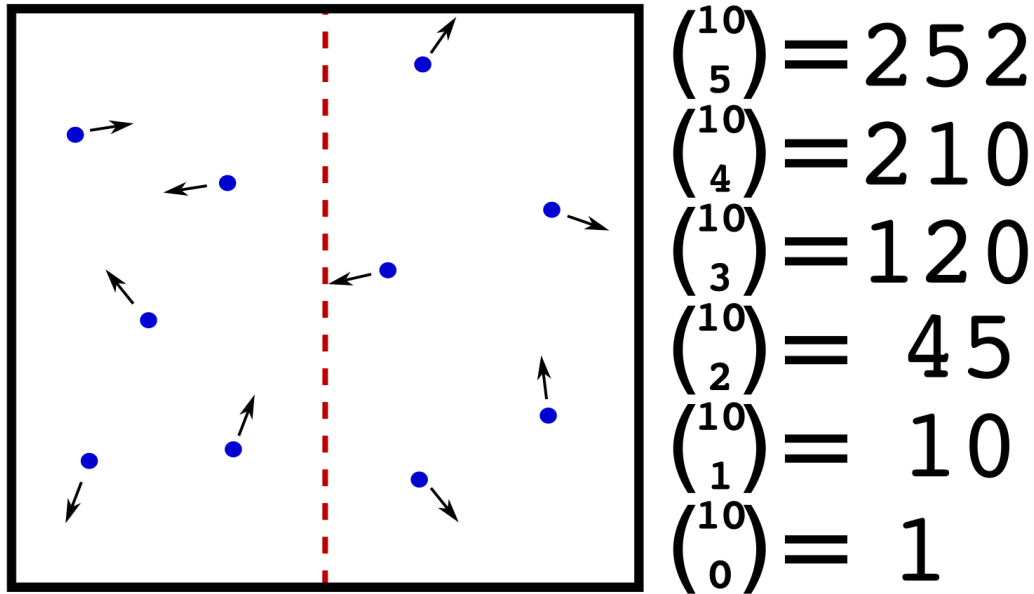
Opierając się o powyższy postulat, możemy w następujący sposób zdefiniować prawdopodobieństwo p_i wystąpienia stanu i (rozumianego jako arbitralny podzbiór mikrostanów w ramach zadanego makrostanu):

$$p_i = \frac{w_i}{W} = \frac{w_i}{\sum w_i} \quad (3.2)$$

gdzie w_i odpowiada liczbie mikrostanów realizujących zadany stan, zaś W całkowitej liczbie stanów realizujących zadany makrostan.

By zobrazować pojęcie entropii, posłużę się przykładem dwuwymiarowego pudła zawierającego N cząsteczek gazu doskonałego (Rys. 3.1 dla 10 cząsteczek). Liczbę mikrostanów odpowiadających danemu podzieleniu cząsteczek pomiędzy prawą i lewą połowę pudła tak, by z wybranej strony znalazło się i cząstek, możemy wyznaczyć przy pomocy kombinacji bez powtórzeń:

$$w_i = \binom{N}{i} = \frac{N!}{i!(N-i)!} \quad (3.3)$$



RYSUNEK 3.1: Schematyczne przedstawienie dwuwymiarowego pudła zawierającego 10 cząsteczek gazu doskonałego oraz liczby mikrostanów odpowiadające kolejnym sposobom rozłożenia cząsteczek pomiędzy dwie strony pudła, wyznaczonym na podstawie równania 3.3.

Zgodnie z intuicją, największe prawdopodobieństwa odpowiadają stanom, w których cząsteczki gazu rozmieszczone są równomiernie pomiędzy części obie pudła. Stąd stan o największej entropii możemy uznać za najbardziej reprezentatywny dla makrostanu o zadanych parametrach N, V i E .

Podobnie mogliśmy rozpatrzyć zespół złożony z W układów realizujących zadany makrostan na który składa się j stanów, z których każdemu odpowiada w_j mikrostanów:

$$\Omega = \frac{W!}{w_1!w_2!\dots w_j!} \quad (3.4)$$

Korzystając ze wzoru Stirlinga oraz definicji entropii równ. 3.1, możemy dalej wyznaczyć wartość entropii dla zespołu W układów:

$$\begin{aligned} S_z &= k_B \ln \Omega = k_B \ln \frac{W!}{w_1!w_2!\dots w_j!} \approx k_B (W \ln W - W - \sum w_i \ln w_i + \sum w_i) \\ &= k_B (\sum w_i \ln W - \sum w_i \ln w_i) = -k_B (\sum w_i \ln \frac{w_i}{W}) \end{aligned} \quad (3.5)$$

Powyższe rozwiązanie w połączeniu z równ. 3.2 pozwala przedstawić entropię pojedynczego układu jako:

$$S = \frac{S_z}{W} = -k_B \sum \frac{w_i}{W} \ln \frac{w_i}{W} = -k_B \sum p_i \ln p_i \quad (3.6)$$

Otrzymane powyżej rozwiązanie może posłużyć do wyznaczenia rozkładów prawdopodobieństwa w innych zespołach statystycznych przy założeniu maksymalizacji entropii zgodnie z drugą zasadą termodynamiki.

Sytuacji w której układ, przy zachowaniu stałej objętości oraz liczby cząstek, może wymieniać energię z otoczeniem odpowiada zespół kanoniczny (N, V, T) . Warunek równowagi termicznej z otoczeniem możemy zapisać w następujący sposób:

$$E = \sum \varepsilon_i p_i \quad (3.7)$$

gdzie przez ε_i oznaczana jest energia danego stanu, przez p_i prawdopodobieństwo wystąpienia tego stanu, zaś przez E średnią energię układu. Korzystając z równań 3.6, 3.7 oraz 3.2 możemy dalej zapisać równanie funkcji pomocniczej pozwalającej na wyznaczenie prawdopodobieństwa wystąpienia stanów w zespole kanonicznym:

$$\Phi(p_i) = -k_B \sum p_i \ln p_i + k_B \alpha (1 - \sum p_i) + k_B \beta (E - \sum \varepsilon_i p_i) \quad (3.8)$$

gdzie α i β są mnożnikami Lagrange'a odpowiadającymi za zachowanie $\sum p_i = 1$ oraz $E = \sum p_i \varepsilon_i$ w trakcie optymalizacji. Ekstremum funkcji charakteryzowane jest przez zerowanie się jej pierwszych pochodnych cząstkowych:

$$0 = \frac{\partial \Phi(p_i)}{\partial p_i} = -k_B (p_i + 1) - k_B \alpha - \sum -k_B \beta \varepsilon_i \quad (3.9)$$

co pozwala zapisać:

$$p_i = \exp(-\alpha - 1) \exp(-\beta \varepsilon_i) \quad (3.10)$$

człon $\exp(-\alpha - 1)$ jest stały i można go potraktować jako stałą normalizacyjną, zaś $\beta = \frac{1}{k_B T}$ charakteryzuje otoczenie, z którym układ wymienia ciepło. Pozwala to finalnie zapisać:

$$p_i = \frac{\exp(-\beta \varepsilon_i)}{\sum \exp(-\beta \varepsilon_i)} = \frac{1}{Q(N, V, T)} \exp(-\beta \varepsilon_i) \quad (3.11)$$

gdzie $\frac{1}{Q(N, V, T)}$ jest funkcją rozkładu dla zespołu kanonicznego. Znajomość prawdopodobieństw poszczególnych mikrostanów pozwala na wyznaczenie wartości średniej dla dowolnej obserwabli:

$$\bar{B} = \sum p_i B_i \quad (3.12)$$

Powyższe równanie pozwala pokazać, że funkcja charakterystyczna zespołu kanonicznego odpowiada energii swobodnej układu:



$$\begin{aligned}
A &= U - TS = \sum \varepsilon_i \frac{\exp(-\beta\varepsilon_i)}{Q} - T \left(-k_B \sum \frac{\exp(-\beta\varepsilon_i)}{Q} \ln \frac{\exp(-\beta\varepsilon_i)}{Q} \right) \\
&= \frac{1}{Q} \sum \left(\varepsilon_i \exp(-\beta\varepsilon_i) + k_B T \exp(-\beta\varepsilon_i) (-\beta\varepsilon_i + \ln \frac{1}{Q}) \right) \\
&= \frac{1}{Q} \sum \left(\varepsilon_i \exp(-\beta\varepsilon_i) + k_B T \exp(-\beta\varepsilon_i) \left(-\frac{\varepsilon_i}{k_B T} + \ln \frac{1}{Q} \right) \right) \\
&= \frac{1}{Q} k_B T \ln \frac{1}{Q} \sum \exp(-\beta\varepsilon_i) = -k_B T \ln Q
\end{aligned} \tag{3.13}$$

Wyprowadzenie rozkładu prawdopodobieństwa dla innych zespołów przeprowadza się analogicznie jak dla zespołu (N, V, T) .

Powyższe równania, choć dające solidne podwaliny teoretyczne pod opis układów wielociałowych, rzadko kiedy są bezpośrednio stosowane ze względu na trudności związane z wyznaczeniem prawdopodobieństw wystąpienia wszystkich mikrostanów. Z pomocą przychodzi tutaj hipoteza ergodyczna, która mówi, że uśrednianie wartości wyznaczanego parametru po zespole statystycznym B_w może zostać zastąpione przez uśrednianie wartości tego samego parametru po czasie B_τ :

$$\bar{B}_w = \lim_{\tau \rightarrow \infty} \bar{B}_\tau \tag{3.14}$$

Założenie to przyjmuje, że po nieskończenie długim czasie trajektoria układu powinna przejść przez wszystkie mikrostanu danego makrostanu. W symulacjach komputerowych wartości średnie przybliżane są zwykle w następujący sposób:

$$\bar{B}_\tau = \frac{1}{t} \sum_{j=0}^t B_j \tag{3.15}$$

gdzie przez t oznaczono liczbę kroków czasowych, a przez B_j wartość obserwabli w danym momencie trajektorii j .

3.2 Mechanika kwantowa a pola siłowe

W nierelatywistycznej mechanice kwantowej ewolucję czasową układu opisuje się przy pomocy równania Schrödingera:

$$\hat{H}\Psi(r, t) = i\hbar\Psi(r, t) \tag{3.16}$$

gdzie \hat{H} oznacza operator energii całkowitej Hamiltona, zaś $\Psi(r, t)$ funkcję falową układu zależną od położenia cząstek r oraz czasu t . W sytuacjach, w których operator Hamiltona nie zależy w sposób jawny od czasu, powyższe zagadnienie sprowadza się do wyznaczenia wartości własnych, odpowiadających energii układu E_i , oraz odpowiadających im funkcji własnych Ψ_i , charakteryzujących stan stacjonarny układu:

$$\hat{H}\Psi_i(r) = E_i\Psi_i(r) \quad (3.17)$$

Niestety, powyższe równanie daje się rozwiązać analitycznie w tylko nielicznych przypadkach, jak np. cząstka w pudle potencjału, oscylator harmoniczny czy atom wodoropodobny. Dla bardziej skomplikowanych układów, takich jak cząsteczki, koniecznym jest przyjęcie przynajmniej kilku uproszczeń. Pierwszym z nich jest tzw. przybliżenie adiabatyczne: przyjmuje ono, że elektrony będące ponad 1800 razy lżejsze od najmniejszego jądra atomowego są w stanie w nieskończenie szybki sposób dostosować się do zmiany położeń jąder. Przybliżenie to pozwala na rozseparowanie funkcji falowej całego układu na iloczyn funkcji falowych jąder $f(R)$ oraz elektronów $\psi(r, R)$:

$$\Psi_i(R, r) \approx \psi_i(r; R)f(R) \quad (3.18)$$

w równaniu tym funkcja falowa elektronów $\psi(r; R)$ zależy jedynie parametrycznie od położeń jąder R . Powyższy zabieg umożliwia zapisanie elektronowego równania Schrödingera:

$$\hat{H}_{el}(R)\psi_i(r; R) = E_i^{el}(R)\psi_i(r; R) \quad (3.19)$$

którego rozwiązanie, po zastosowaniu przybliżenia Borna-Oppenheimera zaniedbującego tzw. poprawkę diagonalną, umożliwia zapisanie równania Schrödingera dla jąder poruszających się w potencjale E_i^{el} otrzymanym w równ. 3.19:

$$[T_{nu} + E_i^{el}(R)]f_i(R) = E_i(R)f_i(R) \quad (3.20)$$

Korzystając dalej z twierdzenia Hellmana-Feynmana, mówiącego że siły działające na jądra dane są przez ich ujemny gradient energii potencjalnej:

$$F_k = -\frac{\partial}{\partial \mathbf{R}_k} E_i^{el}(R_1, R_2, \dots, R_N), \quad k = 1, 2, \dots, N \quad (3.21)$$

otrzymujemy siły działające na jądra w badanym układzie, przy pomocy których możemy dalej propagować ruch jąder w klasyczny sposób. Takie podejście stosowane jest w wariacie Borna-Oppenheimera kwantowych symulacji dynamiki molekularnej. [149] Metoda ta pozwala np. na śledzenie reakcji chemicznych w stosunkowo małych układach. [150] Z racji na to, że główny wysiłek obliczeniowy związany jest z wyznaczeniem potencjału w którym poruszają się jądra, potencjał ten przybliża się za pomocą pola siłowego. Umożliwia to zarówno znaczące wydłużenie przeprowadzanych symulacji (w wariacie pełnoatomowym w skrajnych przypadkach czasu nawet ponad 1 ms [151]) oraz rozmiarów symulowanego układu (nawet 64 mln atomów, jak w przypadku kapsydu HIV-1 [152]).



3.3 Pole siłowe

W klasycznych symulacjach dynamiki molekularnej potencjał działający na jądra przybliża się przy pomocy prostych, addytywnych wkładów, możliwie wiernie oddających zachowanie rzeczywistego układu. Wkłady te skonstruowane są w oparciu proste modele matematyczne, takie jak np. prawo Hooke'a bądź potencjał Lennarda–Jonesa. Postać pola siłowego $V(R)$ wyraża się w następujący sposób:

$$V(R) = V_{bond} + V_{ang} + V_{tors} + V_{impr} + V_{elec} + V_{vdW} \quad (3.22)$$

gdzie poszczególne wkłady opisują energię potencjalną:

V_{bond} – odkształceń długości wiązań

V_{ang} – odkształceń kątów płaskich

V_{tors} – odkształceń kątów torsyjnych

V_{impr} – odkształceń niewłaściwych kątów torsyjnych

V_{elec} – oddziaływań elektrostatycznych

V_{vdW} – oddziaływań dyspersyjnych

Pierwsze cztery składowe pola nazywane są członami wiążącymi, gdyż odpowiadają za modelowanie energii potencjalnej oddziaływań atomów bezpośrednio związanych ze sobą kowalencyjnie. Do ich opisu zwykle stosuje się potencjały dwu-, trój- i czterociałowe, które odpowiadają kolejno długości wiązań, kątów płaskich oraz kątów torsyjnych bądź nieprawidłowych kątów torsyjnych. Kolejne dwa człony opisują oddziaływania pomiędzy atomami, które nie tworzą bezpośrednio wiązań kowalencyjnych.

Zmiany energii potencjalnej układu pod wpływem odkształcenia długości wiązań bądź kątów płaskich zwykle opisuje się przy pomocy potencjału harmonicznego:

$$\begin{aligned} V_{bond} &= \sum \frac{1}{2} k_{gh}^b (r_{gh} - r_{gh}^0)^2 \\ V_{ang} &= \sum \frac{1}{2} k_{ghi}^a (\theta_{ghi} - \theta_{ghi}^0)^2 \end{aligned} \quad (3.23)$$

gdzie przez k oznaczono wartość stałej siłowej, przez r_{gh} odległość między atomami g i h , przez θ_{ghi} wartość kąta płaskiego między atomami g , h oraz i , zaś przez górny indeks 0 oznaczono wartość równowagową. Potencjał harmoniczny dobrze oddaje zachowanie długości wiązań oraz wartości kątów w pobliżu ich minimum energii potencjalnej. Przeszacowuje jednak wartości dla większych odchyłeń od wartości równowagowej, przez co m. in. uniemożliwia zrywanie wiązań.

Kolejny wkład opisuje zmiany energii potencjalnej związane z obrotem wokół wiązań:

$$V_{impr} = \sum \frac{1}{2} k_{ghij}^d (1 + \cos(n\phi_{ghij} - \phi_{ghij}^0)) \quad (3.24)$$



gdzie k jest stałą siłową, ϕ_{ghij} oznacza wartość kąta torsyjnego między atomami g , h , i oraz j , n liczbę minimów przypadających na pełny obrót wokół wiązania, zaś ϕ_{ghij}^0 jest opcjonalnym czynnikiem fazowym. Ponieważ krzywa energii potencjalnej względem obrotu wokół wiązania może wykazywać minima o różnej głębokości, dla dokładniejszego opisu obrotu często stosuje się kombinację kilku członów o różnych wartościach n .

Ostatnim z członów wiążących jest potencjał opisujący tzw. niewłaściwy kąt torsyjny:

$$V_{impr} = \sum \frac{1}{2} k_{ghij}^i (\zeta_{ghij} - \zeta_{ghij}^0)^2 \quad (3.25)$$

gdzie k jest stałą siłową, ζ_{ghij} oznacza wartość niewłaściwego kąta torsyjnego między atomami g , h , i oraz j , zaś ζ_{ghij}^0 wartość równowagową. Stosowanie tego potencjału zwykle ma na celu zachowanie planarności układów takich, jak np. atomy o hybrydyzacji sp^2 bądź hybrydyzacji sp^3 zawierających wolną parę elektronową.

Oddziaływania dyspersyjne między niezwiązanymi ze sobą atomami modeluje się zwykle przy pomocy potencjału Lennarda–Jonesa:

$$V_{vdW} = \sum_i \sum_{j>i} \varepsilon_{ij} \left(\left(\frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^0}{r_{ij}} \right)^6 \right) \quad (3.26)$$

gdzie wartość ε_{ij} odpowiada za głębokość minimum energii potencjalnej, r_{ij} oznacza odległość między atomami i oraz j , zaś górny indeks r_{ij}^0 oznacza odległość atomów w minimum energii potencjalnej. Z racji tego, że energia oddziaływań w potencjale Lennarda–Jonesa zanika z 6 potęgą, wkłady do energii potencjalnej układu pochodzące od atomów znajdujących się w dużych odległościach od siebie są marginalne i można je pominąć. Pozwala to na ograniczenie złożoności obliczeniowej, a dzięki temu na przyspieszenie obliczeń.

Drugi z wkładów niewiązących opisuje oddziaływania elektrostatyczne w oparciu o prawo Coulomba:

$$V_{elec} = \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi \varepsilon_0 \varepsilon_r r_{ij}} \quad (3.27)$$

gdzie q_i oraz q_j oznaczają ładunki cząstkowe atomu i oraz j , ε_0 przenikalność dielektryczną próżni, ε_r przenikalność dielektryczną ośrodka, zaś r_{ij} odległość między atomami i oraz j . Same ładunki cząstkowe atomów są sztucznym konstruktem – nie mają ściśle odpowiadających im obserwabli fizycznych. Dopasowuje się je tak, by oddawały potencjał elektrostatyczny w otoczeniu cząsteczki. Oddziaływania elektrostatyczne zanikają bardzo powoli (z pierwszą potęgą r), co wymaga uwzględnienia wkładów do energii potencjalnej pochodzących od atomów znajdujących się w znacznych odległościach od siebie. W praktyce trudność tę obchodzi się w następujący sposób: dla atomów znajdujących się w bezpośrednim sąsiedztwie (w obrębie tzw. promienia obcięcia) wyznacza się energię oddziaływań bezpośrednio, zaś dla atomów występujących w większych odległościach stosuje się metodę sum Ewalda. [153] Pozwala to na zmniejszenie złożoności obliczeniowej, a dzięki temu na przyspieszenie przeprowadzanych obliczeń.

3.4 Metody wzmocnionego próbkowania

Pomimo daleko idących uproszczeń w ramach opisu układu przy pomocy pól siłowych, obecne moce obliczeniowe komputerów zwykle nie pozwalają na uzyskanie trajektorii o długości umożliwiającej choć przybliżone odtworzenie funkcji rozkładu przy pomocy konwencjonalnych symulacji. Z tej przyczyny opracowano szereg metod tzw. wzmocnionego próbkowania (ang. enhanced sampling), które ułatwiają pokonywanie barier energii swobodnej, a dzięki temu pozwalają na bardziej precyzyjne odtworzenie stanu równowagowego układu.

3.4.1 Metody replik temperaturowych

Symulacje replik temperaturowych opierają się o równoległe prowadzenie symulacji układu w zestawie różnych temperatur $(T_0, T_1, T_2, \dots, T_N)$ przez określony okres czasu, gdzie temperatury są uporządkowane od najniższej T_0 do najwyższej T_N . Po tym okresie czasu podejmowana jest próba wymiany konfiguracji pomiędzy sąsiednimi replikami:

$$\begin{aligned} (X_m, E_m(X_m), T_m) &\rightarrow (X_n, E_m(X_n), T_m) \\ (X_n, E_n(X_n), T_n) &\rightarrow (X_m, E_n(X_m), T_n) \end{aligned} \quad (3.28)$$

gdzie X_m oznacza konfigurację, $E_m(X_m)$ energię potencjalną przy konfiguracji X_m , zaś T_m temperaturę układu w replice m jeszcze przed zajściem wymiany. Układy między sobą mogą różnić się nie tylko temperaturą, ale również funkcjami energii potencjalnej. Równowagowe prawdopodobieństwo wystąpienia takiego stanu w replice temperaturowej m można zapisać jako:

$$P_m = \frac{1}{Q_m} \exp(-\beta_m E_m(X_m)) \quad (3.29)$$

gdzie $\beta_m = \frac{1}{k_B T_m}$, zaś Q_m to konfiguracyjna funkcja podziału dla repliki m . Wymiana między replikami jest akceptowana z zachowaniem warunku równowagi szczegółowej:

$$P_m(X_m), P_n(X_n), W(i \rightarrow j) = P_m(X_n), P_n(X_m), W(i \leftarrow j) \quad (3.30)$$

gdzie $W(i \rightarrow j)$ odpowiada wymianie przedstawionej w równ. 3.28, zaś $W(i \leftarrow j)$ wymianie w przeciwnym kierunku. Pozwala to na wyznaczenie proporcji dla zajścia wymiany między replikami w następujący sposób:

$$\frac{W(i \rightarrow j)}{W(i \leftarrow j)} = \exp(-\Delta_{nm}) \quad (3.31)$$

gdzie:

$$\Delta_{nm} = -\beta_m (E_m(X_n) - E_m(X_m)) - \beta_n (E_n(X_m) - E_n(X_n)) \quad (3.32)$$

dzięki czemu, stosując kryterium Metropolisa, możemy finalnie zapisać szanse na zajście wymiany między replikami:

$$W(i \rightarrow j) = \begin{cases} 1, & \Delta_{nm} \leq 0 \\ \exp(-\Delta_{nm}), & \Delta_{nm} > 0 \end{cases} \quad (3.33)$$

Następnie proces wymian między replikami powtarza się co zadany wcześniej okres czasu. Najwyższą temperaturę T_N dobiera się tak, by układ w łatwy sposób przecho- dził między barierami energii potencjalnej. Jako, że ostatecznie układy niskotem- peraturowe powinny wymienić się z wysokotemperaturowymi, układy niskotem- peraturowe będą mogły przeskakiwać między basenami energii potencjalnej, które rozdzielone są przy pomocy wysokich barier.

Jednym z wariantów symulacji replik temperaturowych są repliki temperaturowe z wygrzewaniem substancji rozpuszczonej (ang. replica exchange with solute tem- pering), [154] w których system symulacyjny dzieli się na substancję rozpuszczoną p , np. białko, oraz rozpuszczalnik w , np. wodę. Energię potencjalną tego układu można zapisać w następujący sposób:

$$E_0(X) = E_p(X) + E_{pw}(X) + E_w(X) \quad (3.34)$$

gdzie $E_0(X)$, $E_p(X)$, $E_{pw}(X)$ i $E_w(X)$, to kolejno całkowita energia potencjalna, energia potencjalna oddziaływań wewnętrznych substancji rozpuszczonej, energia oddziaływań między substancją rozpuszczoną a rozpuszczalnikiem oraz energia oddziaływań między cząsteczkami rozpuszczalnika. Takie rozdzielenie całkowi- tej energii potencjalnej na człony pozwala dalej na przeskalowanie każdego z nich osobno dla każdej z replik w następujący sposób:

$$E_m(X) = E_p(X) + \frac{\beta_0 + \beta_m}{2\beta_m} E_{pw}(X) + \frac{\beta_0}{\beta_m} E_w(X) \quad (3.35)$$

gdzie E_m to energia potencjalna repliki m , zaś $\beta_m = \frac{1}{k_B T_m}$. Podstawienie równ. 3.35 do równ. 3.32 pozwala zapisać:

$$\Delta_{nm} = (\beta_n - \beta_m) \left((E_p(X_m) + \frac{1}{2} E_{pw}(X_m)) - (E_p(X_n) + \frac{1}{2} E_{pw}(X_n)) \right) \quad (3.36)$$

Jak można zauważyć, z powyższego równania zniknął człon $E_w(X)$, który stanowił największy z wkładów do energii potencjalnej układu. Powyższe podejście pozwala zatem na istotne zwiększenie prawdopodobieństwa wymian pomiędzy sąsiadują- cymi replikami.

Powyższa metoda, choć bardzo atrakcyjna dla stosunkowo małych układów, nie jest pozbawiona wad: 1. w wielu przypadkach okazuje się nadal niedostatecznie szybka 2. zwykle nie pozwala też na szczegółowy opis stanów przejściowych, jako że zwy- kle pojawiają się one jedynie w wyższych replikach, które są niefizyczne 3. pozwala na opis termodynamiczny układu, nie niesie z sobą jednak informacji o kinetyce badanego procesu (następują przeskoki trajektorii w trakcie symulacji wynikające z wymian). W sytuacjach, w których zainteresowanie skupione jest na konkretnej



zmianie konformacyjnej, wykorzystuje się metody oparte o współrzędne opisujące, które przedstawiłem poniżej.

3.4.2 Metadynamika

W przeciwieństwie do symulacji replik temperaturowych, symulacje metadynamiki skupiają się na reprezentatywnym wypróbkowaniu przestrzeni konformacyjnej wzdłuż wybranej współrzędnej opisującej. Współrzędna ta powinna być wybierana tak, by w uproszczony sposób opisać najwolniejsze ruchy występujące w badanym systemie, takie jak np. otwieranie centrum katalitycznego enzymu czy tworzenie kompleksu przez parę białek. Prowadzone symulacje mają na celu wyznaczenie tzw. potencjału średniej siły wzdłuż wybranej współrzędnej opisującej, który można utożsamiać z profilem energii swobodnej wzdłuż tej współrzędnej. [155] Potencjał średniej siły możemy zapisać w następujący sposób:

$$F = -k_B T \lim_{t \rightarrow \infty} \ln(N(s, t)) \quad (3.37)$$

gdzie $N(s, t)$ jest histogramem zmiennej s uzyskany w ramach nieskończenie długiej, równowagowej symulacji. W symulacjach metadynamiki ewolucja czasowa systemu jest zaburzana przez dodatkowy, zależny od historii trajektorii potencjał, który złożony jest z sumy krzywych Gaussa deponowanych wzdłuż wybranych współrzędnych opisujących [156]:

$$V_G(s, t) = \sum_{t'=0}^t \omega \exp\left(-\sum_{i=1}^d \frac{(S_i(R) - S_i(R(t')))^2}{2\sigma_i^2}\right) \quad (3.38)$$

gdzie ω opisuje szybkość z jaką dodawany jest potencjał, d oznacza liczbę współrzędnych opisujących, $S_i(R)$ jest jedną z funkcji położeń zestawu atomów R , a σ odchyleniem standardowym (nazywanym też szerokością) dodawanej krzywej Gaussa. Szybkość dodawania potencjału można przedstawić w następujący sposób:

$$\omega = \frac{W}{\tau} \quad (3.39)$$

gdzie W oznacza wysokość dodawanego potencjału, zaś τ okres, co który dodawana jest kolejna krzywa Gaussa do potencjału $V(s, t)$.

Metadynamika posiada kilka zasadniczych zalet: 1. pozwala na wypróbkowanie rzadkich zdarzeń dzięki wypychaniu systemu z lokalnego minimum energii swobodnej, 2. umożliwia poszukiwanie nowych ścieżek reakcji, jako że układ wykazuje tendencję do ucieczki z lokalnego minimum przez najniższą barierę energii swobodnej, 3. po odpowiednio długim czasie dodany potencjał stanowi oszacowanie profilu energii swobodnej wzdłuż współrzędnej opisującej:

$$V_G(s, t) = -F(S) + C \quad (3.40)$$

gdzie C jest stałą addytywną. Tutaj należy jednak zaznaczyć, że otrzymany potencjał V_G nie uzbiegnie się z czasem do bezwzględnej wartości profilu energii swobodnej (powiększonego o stałą), a oscyluje naokoło prawidłowej wartości. Jednym ze



sposobów ominięcia powyższej trudności jest skalowanie dodawanego potencjału z czasem, co można zapisać w następujący sposób:

$$W = W_0 \exp\left(-\frac{V(s, t)}{k_B \Delta T}\right) \quad (3.41)$$

gdzie W_0 jest początkową wysokością dodawanego potencjału, zaś ΔT parametrem o wymiarze temperatury. Wraz z postępowaniem symulacji wysokość dodawanego potencjału spada, zaś parametr ΔT opisuje szybkość tego spadku. W tym przypadku potencjał uzbiegnia się do następującego wyrażenia:

$$V_G(s, t \rightarrow \infty) = -\frac{\Delta T}{\Delta T + T} F(S) + C \quad (3.42)$$

W miejsce parametru ΔT często podaje się tzw. czynnik skalujący potencjał (ang. bias factor), który ma następującą postać:

$$\gamma = \frac{\Delta T + T}{T} \quad (3.43)$$

jest on oszacowaniem, ilukrotnie zostaną zmniejszone bariery energii swobodnej przez dodany potencjał.

3.4.3 Umbrella sampling

Umbrella sampling to kolejna z metod polegająca na wzmocnionym próbkowaniu przestrzeni konformacyjnej wzdłuż wybranej współrzędnej opisującej. W ramach tej metody przeprowadza się serię symulacji w obecności dodatkowego, zewnętrznego potencjału $V(\zeta)$, który utrzymuje układ w pobliżu kolejnych wybranych wartości współrzędnej opisującej ζ . Dzięki temu możliwe jest reprezentatywne wypróbkowanie wysokoenergetycznych obszarów przestrzeni konformacyjnej wzdłuż wybranej współrzędnej reakcji. W ramach każdej z przeprowadzonych symulacji otrzymujemy rozkład prawdopodobieństwa występowania układu wzdłuż współrzędnej ζ zaburzony przez przyłożony dodatkowy potencjał $V(\zeta)$:

$$p_b(\zeta) = \frac{\int \delta(\zeta(r) - \zeta) \exp\left(-\frac{U(r)+V(\zeta)}{k_B T}\right)}{\int \exp\left(-\frac{U(r)+V(\zeta)}{k_B T}\right)} \quad (3.44)$$

gdzie $U(r)$ oznacza energię potencjalną układu, zaś p_b prawdopodobieństwo wystąpienia układu przy danej wartości współrzędnej ζ dla symulacji prowadzonych wobec dodatkowego potencjału. Jako, że potencjał $V(\zeta)$ jest stały dla danej wartości współrzędnej ζ , człon $\exp\left(-\frac{V(\zeta)}{k_B T}\right)$ można w liczniku wyłączyć przed całkę. Mnożąc dalej licznik i mianownik przez sumę statystyczną dla układu niezaburzonego przez dodatkowy potencjał, otrzymujemy:



$$\begin{aligned}
p_b(\zeta) &= \frac{\int \delta(\zeta(r) - \zeta) \exp\left(-\frac{U(r)}{k_B T}\right) \exp\left(-\frac{V(\zeta)}{k_B T}\right)}{\int \exp\left(-\frac{U(r)+V(\zeta)}{k_B T}\right)} \times \frac{\int \exp\left(-\frac{U(r)}{k_B T}\right)}{\int \exp\left(-\frac{U(r)}{k_B T}\right)} \\
&= \exp\left(-\frac{V(\zeta)}{k_B T}\right) \times \frac{\int \delta(\zeta(r) - \zeta) \exp\left(-\frac{U(r)}{k_B T}\right)}{\int \exp\left(-\frac{U(r)}{k_B T}\right)} \times \frac{\int \exp\left(-\frac{U(r)}{k_B T}\right)}{\int \exp\left(-\frac{U(r)+V(\zeta)}{k_B T}\right)} \quad (3.45) \\
&= \exp\left(-\frac{V(\zeta)}{k_B T}\right) \times p_u(\zeta) \times \frac{\int \exp\left(-\frac{U(r)}{k_B T}\right)}{\int \exp\left(-\frac{U(r)+V(\zeta)}{k_B T}\right)}
\end{aligned}$$

gdzie przez $p_u(\zeta)$ oznaczono rozkład prawdopodobieństwa dla niezaburzonego systemu. Przekształcając równanie 3.45, otrzymujemy postać na niezaburzony rozkład prawdopodobieństwa $p_u(\zeta)$ względem współrzędnej ζ :

$$p_u(\zeta) = p_b(\zeta) \times \exp\left(\frac{V(\zeta)}{k_B T}\right) \times \frac{\int \exp\left(-\frac{U(r)}{k_B T}\right) \exp\left(\frac{V(\zeta)}{k_B T}\right)}{\int \exp\left(-\frac{U(r)}{k_B T}\right)} \quad (3.46)$$

ostatni z członów równania 3.46 jest równy wartości oczekiwanej dodanego potencjału $\exp\left(\frac{V(\zeta)}{k_B T}\right)$ po rozkładzie niezaburzonym i może być traktowany jako wkład do energii swobodnej układu w wyniku przyłożenia dodatkowego potencjału:

$$F = -k_B T \ln \left\langle \exp\left(-\frac{V(\zeta)}{k_B T}\right) \right\rangle \quad (3.47)$$

Ostatecznie pozwala to zapisać:

$$p_u(\zeta) = p_b(\zeta) \times \exp\left(\frac{V(\zeta) - F}{k_B T}\right) \quad (3.48)$$

wartości parametru F nie można otrzymać bezpośrednio z symulacji. Optymalizuje się je w ramach łączenia wyników z różnych okien, np. metodą ważonych histogramów (ang. weighted histogram analysis method). [157]

3.5 Procedura symulacyjna

Symulacje prowadzone były przy pomocy pakietu Gromacs 5.0.4 [158] w zespole NPT z wykorzystaniem pola siłowego CHARMM36 [159] dla białek i dla lipidów oraz przy zastosowaniu modelu wody TIP3P. [160] Symulowane układy błonowe zostały zbudowane przy pomocy serwera Charmm-Gui Membrane Builder. [161] Dla systemów zawierających białka Hsp bądź domenę 4HB temperatura układu utrzymywana była na poziomie 310 K, zaś w układach zawierających γ -sekretezę na poziomie 320 K przy pomocy termostatu Nose-Hoovera. [162] Ciśnienie było utrzymywane izotropowo dla układów zawierających białka Hsp bądź domenę 4HB oraz semi-izotropowo dla systemów błonowych (oddzielnie dla XY i Z) na poziomie 1

bara przy pomocy algorytmu Parrinello-Rahmana. [163] Ciągłość symulowanego środowiska uzyskano przy zastosowaniu periodycznych warunków brzegowych. Oddziaływania elektrostatyczne wyznaczone są przy pomocy algorytmu PME z promieniem obciążenia wynoszącym 1.2 nm. [153] Oddziaływania dyspersyjne wyznaczone były z zastosowaniem promienia obciążenia wynoszącym 1 nm dla Hsp70 bądź 1.2 nm z zastosowaniem funkcji przełączającej Force-Switch dla odległości powyżej 1 nm dla układów błonowych. Wiązania zawierające atom wodoru usztywniane były przy pomocy algorytmu P-Lincs. [164] Równania ruchu całkowane były metodą żabiego skoku Verleta dla kroku czasowego 2 fs. [165] Do symulacji dynamiki sterowanej, umbrella sampling [166], metadynamiki [167] oraz wymiany replik temperaturowych [154] wykorzystano wtyczkę Plumed 2.2. [168] Wizualizacje wykonano przy pomocy pakietu VMD. [169]

3.5.1 Procedury symulacyjne dla systemów zawierających białka HSP

Relaksacja struktury modelu homologicznego białka Ssq1

Celem zrelaksowania struktury homologicznego modelu białka Ssq1 (przygotowanej przez dr. Bartłomieja Tomiczka) przeprowadziłem symulację o długości 2 μ s układu złożonego z białka Ssq1 ze związanym kompleksem ATP z jodem Mg^{2+} , solwatowanych przez 24354 cząsteczek wody. Dodatkowo dodałem 65 kationów potasu i 79 anionów chlorkowych celem zapewnienia fizjologicznej siły jonowej. Przez pierwsze 200 ns konformacja łańcucha głównego białka była usztywniona, a ATP wraz z jodem magnezu i jego otoczką solwatacyjną były utrzymywane w kieszeni wiążącej ATP celem zrelaksowania łańcuchów bocznych białka. Rozkład potencjału elektrostatycznego wokół białka Ssq1 wyznaczyłem jako średnią z rozwiązań liniowego równania Poissona-Boltzmana dla 100 klatek symulacji równowagowej zapisywanych co 10 ns z ostatniej 1 μ s. Do rozwiązania liniowego równania Poissona-Boltzmana wykorzystałem program APBS [170] oraz wtyczkę PDB2PQR [171].

Spontaniczne wiązanie białka Jac1 do białka Ssq1

Celem sprawdzenia, czy przy pomocy symulacji jesteśmy w stanie zaobserwować zdarzenia samorzutnego wiązania białka Jac1 do białka Ssq1, przeprowadziłem zestaw 5 równowagowych symulacji, w których system złożony był z białka Ssq1 ze związanym kompleksem ATP- Mg^{2+} oraz białka Jac1 odseparowanych od siebie, solwatowanych przez 85113 cząsteczek wody z dodatkowymi 287 jonami sodu oraz 278 jonami chlorkowymi dla zapewnienia fizjologicznej siły jonowej roztworu. Białka były wstępnie preorientowane tak, by łańcuchy główne naładowanych obszarów NBD Ssq1 (reszty 244–255 oraz 425–429) i helisy 2 Jac1 (reszty 34–44) znajdowały się w odległości około 3 nm od siebie. Symulacje zostały przerwane po 60 ns, gdy w 2 z 5 przypadków zaobserwowano spontaniczne utworzenie się kompleksu Jac1-Ssq1.

Symulacje umbrella sampling dla asocjacji białka Jac1 do białka Ssq1

W celu wyznaczenia profilu energii swobodnej asocjacji białka Jac1 do białka Ssq1, wykorzystano metodę hamiltonian replica exchange umbrella sampling.[166] Symulacje służące do wyznaczenia profilu energii swobodnej były prowadzone na podstawie struktur uzyskanych z trajektorii równowagowych symulacji spontanicznego wiązania białka Jac1 do białka Ssq1, opisanych powyżej. Jako współrzędną reakcji opisującą proces wiązania wybrano odległość pomiędzy łańcuchami głównymi naładowanych obszarów NBD Ssq1 (reszty 244–255 oraz 425–429) i helisy 2 Jac1 (reszty 34–44). Okna symulacyjne rozmieszczono w zakresie od 1.2 nm do 1.7 nm co 0.1 nm, a w zakresie od 1.7 nm do 4.2 nm co 0.2 nm. By symulacje prowadzić w sposób efektywny co do wykorzystania zasobów obliczeniowych, obszary wymian podzielono na dwie części – dla zakresów odległości 1.2 nm to 2.4 nm oraz 2.4–4.2 nm. Stałe siłowe dobrano tak, by efektywnie próbować obszar bezpośredniego wiązania (centra potencjału oraz odpowiadające im stałe siłowe przedstawione są w Tab. 3.1). W przypadku symulacji przeprowadzonych dla wiązania mutantu R37A R41A klatki startowe zostały otrzymane przez zamianę reszt argininy na alaninę w klatkach startowych wykorzystanych do przeprowadzenia symulacji umbrella sampling dla białka dzikiego. Symulacje te były prowadzone według tego samego protokołu. Łącznie dla każdego okna otrzymano trajektorie o długości 400 ns, na podstawie których uzyskano profile energii swobodnej metodą ważonych histogramów,[157] zaś niepewności wyznaczono uwzględniając autokorelację uzyskanych przebiegów czasowych współrzędnej.

TABLICA 3.1: Położenia centrów potencjału oraz odpowiadające im stałe siłowe symulacji umbrella sampling dla asocjacji białka Jac1 do Ssq1. Przez c oznaczono położenie centrum potencjału, a przez k wartość stałej siłowej

Bliższe odległości			Dalsze odległości		
l.p.	c [nm]	k [kJ mol ⁻¹ nm ⁻²]	l.p.	c [nm]	k [kJ mol ⁻¹ nm ⁻²]
1	1.2	3000	1	2.4	1000
2	1.3	3500	2	2.6	1000
3	1.4	4000	3	2.8	1000
4	1.5	3000	4	2.4	1000
5	1.6	3000	5	3.2	1000
6	1.7	2500	6	3.4	500
7	1.8	2000	7	3.6	500
8	2.0	1500	8	3.8	500
9	2.2	1000	9	4.0	500
10	2.4	1000	10	4.2	500

Dokowanie molekularne i równowagowe symulacje kompleksu Jac1-Ssq1

W celu uzyskania reprezentatywnych struktur białek Ssq1 z Jac1 do dokowania molekularnego, na początku przeprowadziłem zestaw symulacji równowagowych każdego z tych białek. W celu uzyskania możliwie różnorodnych konformacji obu białek, przeprowadziłem dla nich symulacje równowagowe w dwóch różnych polach siłowych: CHARMM36 o długości odpowiednio 10 μs dla Jac1 i 2 μs dla Ssq1, oraz

AMBER99SB-ILDN o długości 5 μ s dla Jac1 i 2 μ s dla Ssq1. Dla każdej z tych trajektorii przeprowadziłem klastrowanie przy pomocy programu gmx cluster z promieniem obcięcia odpowiednio 0.35 nm dla Jac1 oraz 0.7 nm dla Ssq1, dzięki czemu otrzymałem odpowiednio 12 różnych konformacji białka Jac1 oraz 7 dla Ssq1. Dało to łącznie 84 pary konformerów, które zostały następnie wykorzystane do dokowania molekularnego (przeprowadzonego przez dr. Bartłomeja Tomiczka). W wyniku dokowania otrzymał on 8973 struktur kompleksów, które następnie poddał wstępnej selekcji w oparciu o 4 kryteria: najbardziej korzystna energia wiązania, minimalna odległość helisy 2 Jac1 od NBD, minimalna odległość helisy 3 Jac1 od NBD oraz minimalna odległość domeny J Jac1 od NBD. W kolejnym kroku porównał ze sobą otrzymane sposoby związania a powtarzające się struktury usunięto, dzięki czemu finalnie otrzymał 33 różnorodnych struktur kompleksu Jac1-Ssq1. Dla każdej z tych struktur przeprowadziłem symulacje o długości od 500 ns do 1.5 μ s, które pozwoliły na dodatkowe przeszukanie możliwych stanów związania białka Jac1 do Ssq1. Sumaryczna długość uzyskanych trajektorii wyniosła 25 μ s. Następnie przeprowadziłem klastrowanie tych trajektorii przy pomocy programu gmx cluster stosując algorytm single linkage, w pierwszym kroku dopasowując do siebie położenie fragmentów domeny NBD odpowiedzialnego za wiązanie białek z domeną J, zaś w kolejnym klastrując położenia domeny J, przyjmując promień obcięcia wynoszący 18Å pomiędzy ciężkimi atomami helis 2 i 3 domeny J.

Relaksacja konformacyjna pętli białka Jac1

Symulowany system złożony był z pojedynczej cząsteczki białka Jac1 solwatowanej przez 17241 cząsteczek wody oraz z 50 kationów potasu i 50 anionów chlorkowych. By scharakteryzować możliwe stany konformacyjne pętli HPD białka Jac1, przeprowadzono zestaw 21 symulacji metadynamiki w wariacie well-tempered [172] z wymianą replik (ang. bias-exchange metadynamics). [173] W 20 z tych symulacji potencjał przykładany był do jednego z kątów ϕ bądź ψ reszt 54–63 pętli HPD. Ostatnia symulacja prowadzona była bez przykładania dodatkowego potencjału celem uzyskania trajektorii zawierającej równowagową dystrybucję kątów torsyjnych. Powyższy zestaw symulacji prowadzono przez 400 ns, co daje sumaryczną długość uzyskanych trajektorii wynoszącą 8.4 μ s. Dalej, by scharakteryzować dostępne stany konformacyjne, przeprowadziłem analizę głównych składowych na trajektorii zrzutowanej na przestrzeń kątów torsyjnych pętli przy pomocy pakietu MSMbuilder. [174] Wyznaczone dwa pierwsze wektory własne, odpowiadające dwóm głównym ruchom konformacyjnym pętli, wykorzystano dalej jako współrzędne opisujące mapę energii swobodnej pętli. Dynamika pętli w stanie zwiazanym Jac1 do Ssq1 została porównana przez zmapowanie struktury domeny J z największego klastra otrzymanego w ramach sekcji "Dokowanie molekularne i równowagowe symulacje kompleksu Jac1-Ssq1" wyznaczone wektory własne uzyskane z analizy głównych składowych wolnego białka Jac1.

Symulacje 2D metadynamiki dla oddziaływania Arg207/167 z jej natywnymi kontaktami w białku Ssq1/DnaK w obecności bądź pod nieobecność zwiazanego białka JDP

Mapy energii swobodnej dla zrywania kontaktów kluczowej Arg ze swoimi partnerami na łączniku (Asp429 w Ssq1, Asp393 w DnaK) i SBD- β (Asp517 w Ssq1, Asp481

w DnaK) wyznaczyłem stosując metadynamikę w wariacie well-tempered. [172] Jako współrzędne reakcji wybrano odległość między guanidyną kluczowej z Arg, a każdą z grup karboksylowych tworzących natywny kontakt ze wspomnianą Arg (Asp429 i Asp517 dla Ssq1, Asp393 i Asp481 dla DnaK). By usprawnić próbkowanie powierzchni rozpinanej przez te współrzędne, zastosowałem metodę multiple walkers, [175] wybierając klatki startowe z symulacji równowagowych kompleksów Ssq1-Jac1 oraz z dodatkowo przeprowadzonej równowagowej symulacji DnaK-DnaJ o długości 2 μ s. Dodatkowo, by ograniczyć rozmiar układu celem efektywniejszego wypróbkowania przestrzeni konfiguracyjnej, domena SBD- α została skrócona do fragmentu bezpośrednio oddziałującego z domeną NBD (przez usunięcie reszt 570–657 dla Ssq1 i 535–602 dla DnaK). Czynniki skalujące dodawany potencjał (ang. bias factor) został ustalony na wartość 15, a temperatura odniesienia na 310 K. Potencjał w postaci krzywych Gaussa o szerokości wynoszącej 0.04 nm oraz początkowej wysokości 0.1 kJ mol⁻¹ dodawany był co 10 ps. Sumaryczny czas przeprowadzonych symulacji wyniósł 10 μ s. Otrzymane mapy energii swobodnej były wyznaczone jako średnia z 15 map wyznaczonych co 25 ns z ostatnich 350 ns, by zniwelować fluktuacje energii swobodnej.

Symulacje 1D metadynamiki dla zmiany partnerów wiązania Arg207 w białku Ssq1 oraz Arg167 w DnaK z natywnych na resztę Asp motywu HPD białek JDP

Profile energii swobodnej dla zmiany partnerów wiązania Arg207 (Ssq1) oraz Arg167 (DnaK) z natywnych (Asp429 i Asp517 w Ssq oraz Asp393 oraz Asp481 w DnaK) na resztę asparaginianową motywu HPD wyznaczyłem przy pomocy wariantu well-tempered metadynamiki [172] w kombinacji z metodą multiple walkers. [175] Jako współrzędną reakcji wybrałem różnicę pomiędzy odległością centrum masy guanidyny reszty Arg207/Arg167 od reszt karboksylowych reszt Asp429 i Asp517 bądź Asp393 i Asp481 a odległością pomiędzy Arg207/Arg167 a grupą karboksylową Asp50 białka Jac1 bądź Asp35 DnaJ. Do symulacji wykorzystano te same klatki startowe, co dla symulacji 2D metadynamiki z punktu wcześniejszego. Czynniki skalujące dodawany potencjał (ang. bias factor) został ustalony na wartość 8, a temperatura odniesienia na 310 K. Potencjał w postaci krzywych Gaussa o szerokości wynoszącej 0.04 nm oraz początkowej wysokości 0.06 kJ mol⁻¹ dodawany był co 10 ps. Sumaryczny czas przeprowadzonych symulacji wyniósł 10 μ s. Dodatkowo, by zawęzić badaną przestrzeń konfiguracyjną do stanu, w którym pętla HPD orientuje się w pobliżu kluczowej Arg, odległości pomiędzy resztą Arg207 Ssq1 a każdą z Asp429 łącznika, Asp517 SBD i Asp50 białka Jac1 (analogicznie dla pary DnaK-DnaJ) ograniczyłem do wartości 1.5 nm przy pomocy zewnętrznego potencjału o wartości 3000 kJ mol⁻¹ nm⁻². Finalne profile energii swobodnej otrzymane były jako średnia z 8 profili wyznaczonych co 20 ns z ostatnich 140 ns, w celu zniwelowania fluktuacji energii swobodnej.

Symulacje dynamiki sterowanej dla rozdysocjowania domen NBD i NBD- β białek Hsp70 w stanie wolnym i w kompleksie z białkiem JDP

Jako współrzędną reakcji dla dysocjacji domeny SBD- β od domeny NBD wybrałem odległość pomiędzy centrum masy atomów C α domeny SBD- β a centrum masy atomów C α fragmentów NBD znajdujących się w bezpośrednim kontakcie z SBD- β (reszty 112–143 i 186–262 w przypadku Ssq1 oraz reszty 75–104 oraz 146–226 w

przypadku DnaK). Za wyjątkiem pełnych wariantów kompleksu Hsp70 z białkiem JDP, w przypadku przeprowadzanych symulacji zarówno dynamiki sterowanej, jak również następczych symulacji umbrella sampling (punkt dalej) domena SBD- α została całkowicie usunięta. Dla każdego z rozważanych przypadków (Ssq1/DnaK wolne bądź w kompleksie z białkiem JDP) przeprowadziłem po dwie symulacje dynamiki sterowanej o długości 500 ns: w pierwszej odległość między centrami mas była zwiększana ze stałą prędkością od odległości 2.58 nm do 3.58 nm przy pomocy ściany potencjału o stałej siłowej 2500 kJ mol⁻¹ nm⁻². W drugiej, potencjał o stałej siłowej 2500 kJ mol⁻¹ nm⁻² przykładany był do minimalnej odległości pomiędzy wytypowanymi wyżej grupami atomów NBD i SBD- β , wyrażonej wzorem:

$$d = \frac{\beta}{\sum \exp \frac{\beta}{d_{ij}}} \quad (3.49)$$

gdzie β jest parametrem dobieranym, przyjętym tutaj jako 8, a d_{ij} jest odległością między atomem i NBD oraz atomem j SBD- β . W trakcie symulacji wartość tego parametru zmieniano ze stałą prędkością od 0.4 nm do 1.4 nm. W przypadku analizowanych mutantów (dla Ssq1: Arg207Asp; dla DnaK: Arg167Asp, Arg167Cys Asp481Cys utleniony, kompleks z mutantem Asp35Ala białka DnaJ, z helisą 2 bądź helisą 2 i pętlą HPD białka DnaJ) procedura oddalania od siebie NBD i SBD- β została poprzedzona krótką (200 ns) symulacją równowagową.

Symulacje umbrella sampling dla rozdysocjowania domen NBD i NBD- β białek Hsp70 w stanie wolnym i w kompleksie z białkiem JDP

Do wyznaczenia profili energii swobodnej dla dysocjacji domen NBD i SBD- β w białkach HSP70 zastosowano metodę hamiltonian replica exchange umbrella sampling. Okna wybrano dla zakresu współrzędnej od 2.5 nm do 2.8 nm co 0.1 nm. Dla każdego z okien wygenerowano trajektorię o długości 1 μ s. Profile energii swobodnej otrzymałem metodą ważonych histogramów.[157] Niepewności profili energii swobodnej wyznaczyłem uwzględniając autokorelację przebiegów czasowych współrzędnej. Uzyskane trajektorie zostały wykorzystane do wyznaczenia rozkładów prawdopodobieństwa wzdłuż współrzędnych zdefiniowanych jako odległość między centrami mas każdej z subdomen IB, IA i IIA NBD a centrum masy SBD- β bądź jako kąt torsyjny między subdomenami IB-IA-IIA-IIB NBD przeważonych zgodnie z wagami $\exp(V_i(r) - F_i)/k_B T$, gdzie $V_i(r)$ oznacza wartość przyłożonego potencjału w danej klatce, zaś F_i energię swobodną dodaną do i -tego okna symulacji w wyniku przyłożenia dodatkowego potencjału.

Analiza głównych składowych dla zmian w obrębie NBD białka DnaK pod wpływem wymiany nukleotydu bądź związania domeny J białka DnaJ

Wektory własne dla zmiany konformacyjnej NBD pod wpływem wymiany nukleotydu wyznaczyłem na podstawie symulacji pełnego wariantu białka DnaK w stanie otwartym [113] oraz zamkniętym [115] o długości 2 μ s każda. Do wyznaczenia wektorów własnych z uzyskanych trajektorii wybrałem łańcuch główny NBD (reszty 4–386). Wyznaczenie macierzy kowariancji, jej diagonalizacja oraz zrzutowanie trajektorii na wektory własne przeprowadzono przy pomocy pakietu Gromacs (gmx covar i gmx ana eig). [158] Następnie trajektoria o długości 2 μ s pełnego wariantu



DnaK w kompleksie z domeną J białka DnaJ była zrzutowana na otrzymany wyżej pierwszy wyktor własny.

3.5.2 Budowa symulowanych układów dla systemów zawierających domenę 4HB

Równowagowe symulacje domeny 4HB oraz wyznaczanie wkładów do rdzenia hydrofobowego domeny

Symulowane układy złożone były z pojedynczej kopii domeny 4HB solwatowanej przez ok. 7000 cząsteczek wody, z dodatkowo dodanymi kationami sodu i anionami chlorkowymi do zapewnienia fizjologicznej siły jonowej. Struktury domen wzięte były albo z bazy PDB (*S. cerevisiae*, [98] *C. thermophilum*, [125] ludzka [131]) bądź przewidziane metodą modelowania homologicznego przeprowadzonego przez dr. Bartłomieja Tomiczka. Do wyznaczenia powierzchni fragmentów hydrofobowych domeny 4HB współtworzących hydrofobowy rdzeń domeny, przeprowadziłem równowagowe symulacje dynamiki molekularnej o długości 1 μ s dla struktur z bazy PDB bądź 1.8 μ s dla modeli homologicznych, dla których do analiz brałem ostatnie 1 μ s trajektorii, traktując pierwsze 800 ns wyłącznie jako relaksację modelu. Powierzchnie reszt hydrofobowych fragmentów domeny 4HB składające się na rdzeń białka wyznaczono przy pomocy programu g_sas pakietu Gromacs.[176] Wkłady do całkowitej powierzchni hydrofobowej rdzenia wyznaczyłem jako różnicę między całkowitą powierzchnią reszt hydrofobowych wybranych fragmentów białka (patrz Tab. 3.2 i Tab. 3.3) a ich powierzchnią wyeksponowaną do wody.

TABLICA 3.2: Definicje fragmentów domeny 4HB dla struktur z bazy PDB.

	Zakresy reszt aminokwasowych		
	<i>C. thermophilum</i>	<i>S. cerevisiae</i>	human
Helisa 1	355–377	346–365	1–23
Helisa 2	378–398	366–388	24–44
Helisa 3	399–414	389–404	45–60
Helisa 4	415–435	405–422	61–78
C-term	436–446	423–433	79–89

TABLICA 3.3: Definicje fragmentów domeny 4HB dla modeli homologicznych.

	Zakresy reszt aminokwasowych				
	<i>S. cerevisiae</i>	AncA	AncSC	<i>C. albicans</i>	AncP
Helisa 1	1–21	1–26	1–24	1–22	1–20
Helisa 2	22–42	27–46	25–46	23–42	21–42
Helisa 3	43–56	47–61	47–61	43–56	43–57
Helisa 4	57–76	62–82	62–83	57–75	58–74
C-term	77–87	83–93	84–94	76–86	76–86

Rozwijanie domeny 4HB przy pomocy symulacji dynamiki sterowanej

Wkłady C-końca do stabilności domen 4HB były oszacowane jako różnica między pracą wymaganą do mechanicznego rozwinięcia domeny 4HB o pełnej długości a pracą wymaganą do rozwinięcia domeny skróconej o 11 terminalnych reszt (ΔC -term). Wartość pracy koniecznej do rozwinięcia domen oszacowano przy pomocy symulacji dynamiki sterowanej, w której odległość między centrami mas łańcucha głównego Helis I–II oraz III–IV (współrzędna opisująca rozwijanie domeny) była zwiększana przy pomocy zewnętrznego harmonicznego potencjału. Początkowe konfiguracje dla symulacji dynamiki sterowanej brane były z ostatnich 200 ns symulacji równowagowych pełnych wariantów domeny 4HB, zaś warianty ΔC -term otrzymałem przez usunięcie 11 reszt C-końcowych oraz następczą minimalizację energii układu. Dla każdego z układów przeprowadziłem sześć niezależnych symulacji rozwijania białka o długości 500 ns każda, w trakcie których centrum potencjału o stałej siłowej $2500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ było przesuwane ze stałą prędkością 0.04 nm s^{-1} a wartość przyłożonej siły zapisywana była co 10 ps. Otrzymane krzywe przyłożonej siły względem rozciągnięcia domen wzdłuż współrzędnej opisującej rozwijanie domeny otrzymałem przez uśrednienie zmierzonych sił otrzymanych w każdej z 6 niezależnych symulacji dla przedziałów o szerokości 0.08 nm względem współrzędnej opisującej. Niepewności przedstawione dla krzywych przyłożonej siły względem rozciągnięcia wyznaczyłem jako błąd standardowy średniej. Pracę rozwijania domeny wyznaczyłem całkując krzywe przyłożonej siły wzdłuż współrzędnej rozwijania domeny.

Symulacje replik temperaturowych domeny 4HB *S. cerevisiae*

Symulacje replik temperaturowych domeny 4HB *S. cerevisiae* przeprowadziłem w wariacie *solute tempering*, gdzie jako element wygrzewany układu wybrałem całą domenę 4HB. Symulacje prowadziłem w 11 replikach, które obejmowały zakres temperatur od 310 K do 466 K tak, by szansa na wymianę między sąsiadującymi oknami wynosiła 20% . Topologie symulowanych układów przygotowałem przy pomocy skryptu *solute_tempering.py*, dostępnego w repozytorium gitlab naszego zespołu (<https://gitlab.com/KomBioMol>). Dla każdej z replik uzyskałem trajektorie o długości $1.75 \mu\text{s}$, co daje łącznie ponad $19 \mu\text{s}$ symulacji wzmocnionego próbkowania przestrzeni konformacyjnej białka. Stany konformacyjne domeny 4HB zidentyfikowałem, przeprowadzając klastrowanie struktur łańcucha głównego białka uzyskanych dla najniższej repliki temperaturowej przy pomocy programu *gmx cluster* pakietu *Gromacs*, stosując promień obcięcia równy 0.25 nm .

Wyznaczanie energii swobodnej wypętlenia C-końca domeny 4HB *S. cerevisiae*

Profile energii swobodnej dla wypętlenia C-terminalnego fragmentu domeny 4HB *S. cerevisiae* od pełnego fragmentu helikalnego domeny bądź fragmentu pozbawionego helisy 1 (reszt 346–365) wyznaczyłem przy pomocy symulacji metadynamiki w wariacie *well-tempered*. [172] Jako współrzędną opisującą proces wybrałem odległość między centrum masy łańcucha głównego 6 C-terminalnych reszt domeny a centrum masy łańcucha głównego helisy 2 i 4 (reszty 374–387 oraz 406–420). Jako wartość czynnika skalujący potencjał przyjąłem 6, zaś jako temperaturę odniesienia 310 K . Potencjał w postaci krzywych Gaussa o szerokości 0.04 nm oraz początkowej

wysokości 0.04 kJ mol^{-1} dodawany był co 10 ps. W obu przypadkach przeprowadziłem symulacje o długości 1 μs .

3.5.3 Budowa symulowanych układów dla systemów zawierających γ -sekretazę

Symulacje dynamiki sterowanej dla podejścia substratu do wstępnego miejsca wiązania γ -sekretazy

Układy symulacyjne składały się z części transbłonowej γ -sekretazy (usunięto fragment zewnątrzkomórkowy nikastryny złożony z reszt 1–662) oraz domeny transbłonowej (reszty 695–728) białka βCTF , umieszczonych w błonie złożonej z 192 cząsteczek DPPC i 128 cząsteczek cholesterolu jako modelu tratwy lipidowej bądź 192 cząsteczek wody dla czystej błony DPPC, solwatowanych przez 23470 cząsteczek wody, dodatkowo z 65 kationami potasu i 79 anionami chlorkowymi. By uzyskać dłuższe trajektorie, a dzięki temu lepiej uzbieżnione profile energii swobodnej, usunąłem zewnątrzkomórkową domenę nikastryny, która nie wpływa w istotny sposób na dynamikę konformacyjną części transbłonowej enzymu. [177] W ramach prowadzonych symulacji usunąłem również N-terminalny fragment βCTF , który nie jest konieczny dla efektywnego wiązania substratu do γ -sekretazy, [145] zaś jego obecność prowadziła do oddziaływań artefaktowych wynikających z nieobecności domeny zewnątrzkomórkowej nikastryny. Zbliżanie substratu do wstępnego miejsca wiązania γ -sekretazy prowadzone było w ramach symulacji dynamiki sterowanej, dla której jako współrzędną opisującą przyjąłem odległość w płaszczyźnie xy między centrum masy wybranego fragmentu transbłonowego substratu (reszty 706–718) a centrum mas helis 6 i 9 preseniliny (reszty 242–262 oraz 430–451). Dodatkowo, w trakcie symulacji dynamiki sterowanej usztywniłem konformację łańcuchów głównych zarówno substratu jak i helis 6 i 9, przykładając do każdej z nich potencjał na wartość RMSD względem struktury początkowej o wartości $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. Prowadząc symulacje ten sposób w ciągu 200 ns zbliżyłem substrat do miejsca wstępnego wiązania z odległości 2.5 nm do odległości 1 nm przy pomocy zewnętrznego potencjału o wartości $5000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$.

Układ symulacji podejścia substratu do γ -sekretazy w błonie SOPC złożone były z pojedynczej repliki fragmentu błonowego γ -sekretazy oraz białka βCTF osadzonych w błonie składającej się z 274 cząsteczek SOPC solwatowanych przez 19921 cząsteczek wody. Dodatkowo dodano 54 kationów potasu i 68 anionów chlorkowych w celu zapewnienia fizjologicznej siły jonowej. W przypadku symulacji błony SOPC o grubości odpowiadającej modelowi tratwy lipidowej, atomy fosforu każdego z listków błony utrzymywane były w odległości 2.26 nm od centrum masy wszystkich atomów fosforu względem normalnej błony przy pomocy zewnętrznego potencjału o wartości $8000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$.

Symulacje umbrella sampling z wymianą replik dla podejścia substratu w pobliżu centrum katalitycznego

Zakres współrzędnej, względem której prowadziłem symulacje dynamiki sterowanej, podzieliłem na 14 okien rozmieszczonych co 0.1 nm dla odległości między 1 nm a 2.3 nm. Przeprowadzone symulacje umbrella sampling trwały po 1 μs dla układów zawierających błony SOPC oraz 1.5 μs dla układów zawierających błonę DPPC bądź



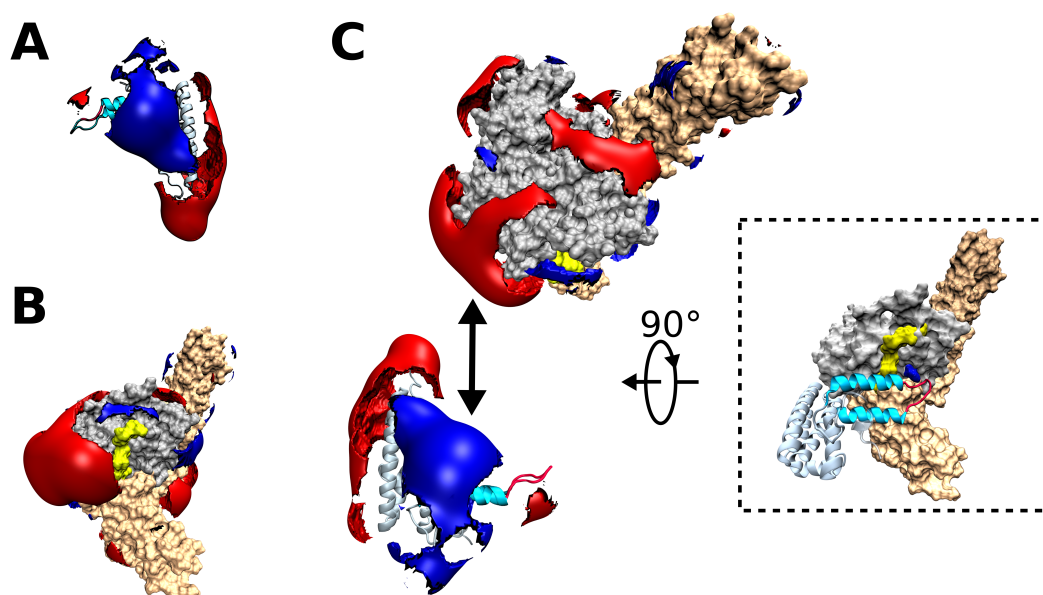
model tratwy. W trakcie tych symulacji odległość między centrami mas utrzymywana była przy pomocy zewnętrznego potencjału o wartości $3500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. Profile energii swobodnej otrzymałem metodą ważonych histogramów.[157] Niepewności profili energii swobodnej wyznaczano uwzględniając autokorelację przebiegów czasowych współrzędnej. Rozkład grubości błon wyznaczono przy pomocy programu g_lomepro. [178]

Rozdział 4

Wyniki i dyskusja

4.1 Oddziaływanie białek HSP70 z białkami pomocniczymi JDP

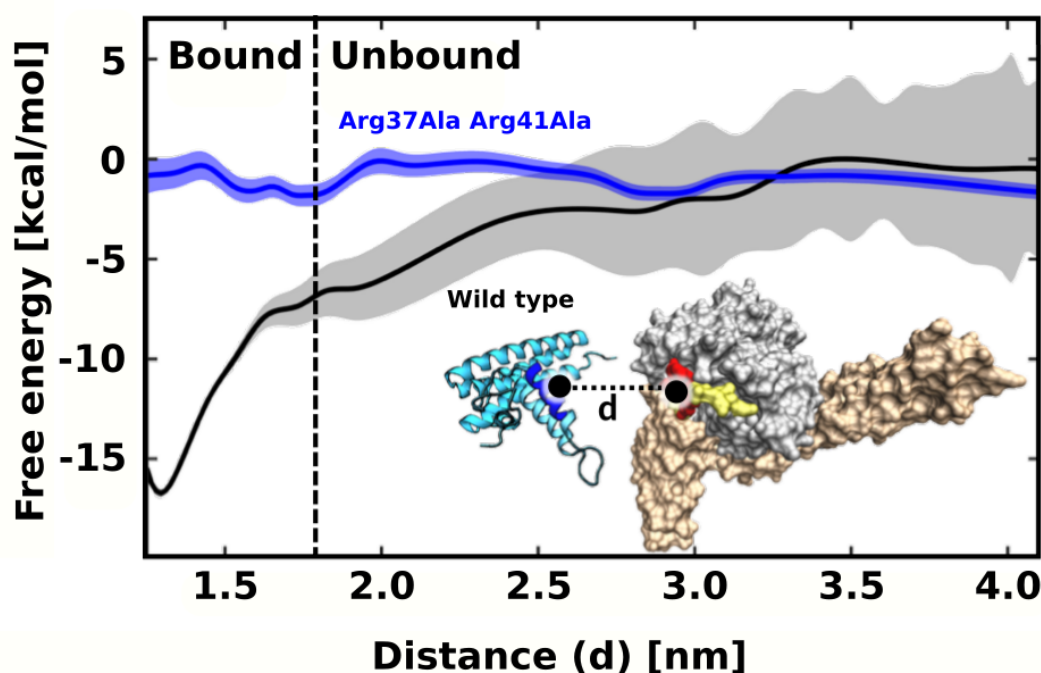
4.1.1 Tworzenie kompleksu Jac1-Ssq1 jest napędzane przez oddziaływania elektrostatyczne



RYSUNEK 4.1: Izopowierzchnie potencjału elektrostatycznego o wartości $2 k_B T$ naokoło białka Jac1 (A) oraz Ssq1 (B). C schematyczne przedstawienie wstępnego ułożenia białek Ssq1 i Jac1 względem siebie w symulacjach spontanicznego tworzenia kompleksu oraz struktura otrzymanego kompleksu.

Jak pokazano w pracy [120], w trakcie koewolucji z białkiem Ssq1, na Jac1 został utworzony dodatnio naładowany obszar w obrębie helisy 2 domeny J (Rys. 4.1). Na tej podstawie założyliśmy, iż w tym samym procesie na Ssq1 powinien pojawić się komplementarny obszar o dużym zgrupowaniu ładunku ujemnego. By sprawdzić tę hipotezę, w pierwszej kolejności przeprowadziłem symulacje równowagowe o długości $2 \mu s$ dla modelu homologicznego białka Ssq1, który otrzymałem od dr. Bartłomieja Tomiczka. Na jej podstawie wyznaczyłem uśredniony rozkład potencjału elektrostatycznego wokół białka Ssq1, który pozwolił na zidentyfikowanie silnie ujemnie naładowanego obszaru zlokalizowanego na β -niciach płata IIA domeny

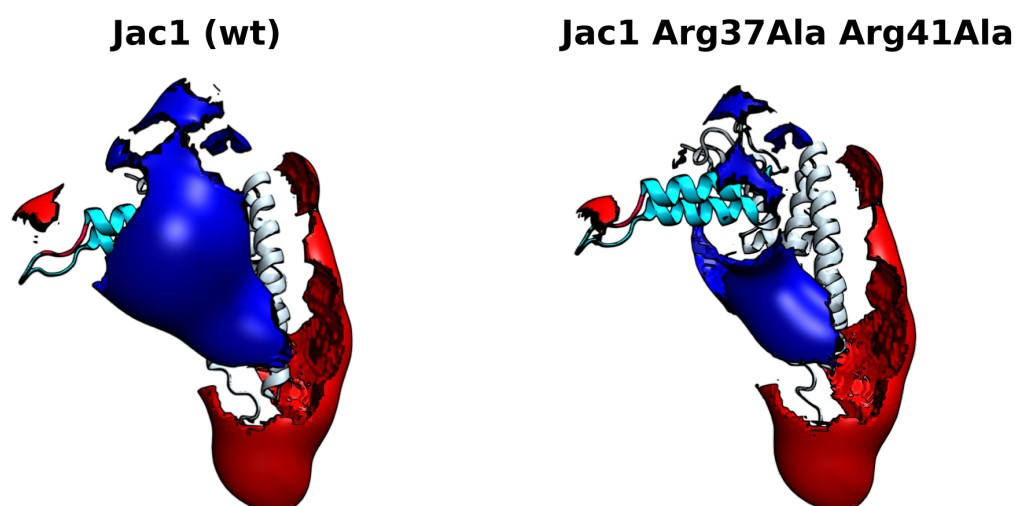
NBD, znajdującego się w bezpośrednim sąsiedztwie powierzchni kontaktu NBD-SBD (Rys. 4.1).



RYSUNEK 4.2: Profile energii swobodnej dla tworzenia kompleksu przez białko Ssq1 z typem dzikim Jac1 (czarny) bądź podwójnym mutantem Jac1 Arg37Ala Arg41Ala (niebieski).

Zachęcony powyższym wynikiem postanowiłem sprawdzić, czy naładowane obszary obserwowane zarówno na Ssq1 jak i Jac1 odpowiadają za rozpoznanie między tymi białkami. W tym celu przeprowadziłem 5 niezależnych symulacji konwencjonalnej dynamiki molekularnej, w których rozseparowane białka ustawiłem względem siebie tak, by naładowane powierzchnie Ssq1 i Jac1, zwrócone ku sobie, znajdowały się w odległości około 3 nm (Rys. 4.1). Symulacje zostały przerwane po ok. 60 ns, gdy w aż dwóch z nich zaobserwowano spontaniczne powstanie kompleksu Ssq1-Jac1. W obu tych przypadkach proces wiązania Jac1 do Ssq1 trwał mniej niż 10 ns i przebiegał najpierw przez przejściowe utworzenie kontaktów między helisą 3 domeny J a domeną SBD- β , po którym helisa 2 białka Jac1 dokowała przez swoje dodatnio naładowane reszty do ujemnie naładowanego obszaru w obrębie β -kartek płatu IIA NBD. W ostatnim kroku pętla zawierająca motyw HPD rozciągała się nad obszarem, w którym znajduje się łącznik domen NBD i SBD, sięgając aż do płata IA NBD. W obu przypadkach kompleks pozostał stabilny od momentu utworzenia do końca symulacji. Szybkość, z jaką Jac1 dokował do miejsca związania na Ssq1 jest zgodna z hipotezą, że za rozpoznanie między tymi białkami odpowiadają dalekozasięgowe oddziaływania elektrostatyczne.

By ilościowo scharakteryzować proces wiązania Jac1 do Ssq1, wyznaczyłem profil energii swobodnej dla tworzenia kompleksu Jac1-Ssq1. W tym celu przeprowadziłem serię symulacji typu umbrella sampling (US), [166] przyjmując jako współrzędną opisującą odległość między centrami mas łańcuchów głównych naładowanych obszarów Jac1 (reszty 34–44) oraz Ssq1 (reszty 244–255 oraz 425–429). Profil



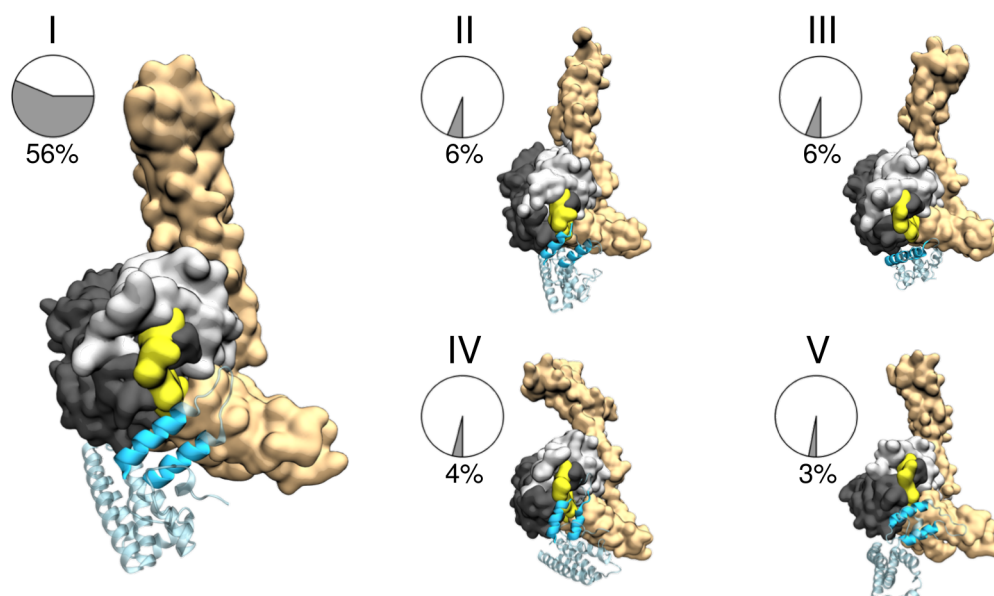
RYSUNEK 4.3: Izopowierzchnie potencjału elektrostatycznego o wartości $2 k_B T$ naokoło szczepu dzikiego białka Jac1 (lewy) oraz podwójnego mutantu Arg37Ala Arg41Ala (prawy)

energii swobodnej tworzenia kompleksu Jac1-Ssq1 (Rys. 4.2) w przypadku typu dzikiego białka Jac1 wykazuje głębokie minimum energii swobodnej dla stanu zasocjowanego białek (wartość współrzędnej opisującej <1.8 nm). Dodatkowo nachylenie profilu energii swobodnej sięga aż do odległości 3.5 nm między domenami, co dodatkowo wskazuje na dominującą rolę oddziaływań elektrostatycznych w tworzeniu kompleksu Jac1-Ssq1. Aby zweryfikować rolę dodatnio naładowanego obszaru w obrębie helisy 2 Jac1 w wiązaniu do sugerowanego miejsca rozpoznania na Ssq1, powtórzyłem obliczenia US dla podwójnego mutantu Jac1 Arg37Ala Arg41Ala. Podwójna mutacja Arg37Ala i Arg41Ala w znaczącym stopniu redukuje wartość potencjału elektrostatycznego na powierzchni helisy 2 Jac1 (Rys. 4.2). Mutant ten nie tworzy też kompleksu z Ssq1, co potwierdzono na drodze doświadczalnej (wyniki nie publikowane). Przeciwnie niż w przypadku typu dzikiego Jac1, dla podwójnego mutantu Jac1 Arg37Ala Arg41Ala na profilu energii swobodnej nie jest obserwowane minimum dla stanu związanego Jac1 z Ssq1 (Rys. 4.2). Wynik ten świadczy o zniesieniu powinowactwa Jac1 do Ssq1 pod wpływem mutacji, co jeszcze raz podkreśla rolę naładowanych reszt w rozpoznaniu i wiązaniu się tych białek.

4.1.2 Kompleks Jac1-Ssq1 wykazuje jeden główny sposób wiązania

By niezależnie zweryfikować uzyskaną strukturę kompleksu Jac1-Ssq1 otrzymaną w wyniku spontanicznego wiązania białek, we współpracy z zespołem prof. Marszałka postanowiłem scharakteryzować możliwe stany wiązania tych białek przy pomocy dokowania molekularnego. W tym celu otrzymałem reprezentatywne struktury izolowanych białek, prowadząc dla nich symulacje w dwóch różnych polach siłowych: w polu CHARMM36 (o długości 10 μ s dla Jac1 i 2 μ s dla Ssq1) oraz AMBER99SB-ILDN (o długości 5 μ s dla Jac1 i 2 μ s dla Ssq1). Dla każdej z uzyskanych trajektorii przeprowadziłem klastrowanie dzięki czemu otrzymałem odpowiednio 12 różnych konformacji białka Jac1 oraz 7 konformacji Ssq1. Dało to 84 różne pary konformerów, na podstawie których dr Bartłomiej Tomiczek wyznaczył

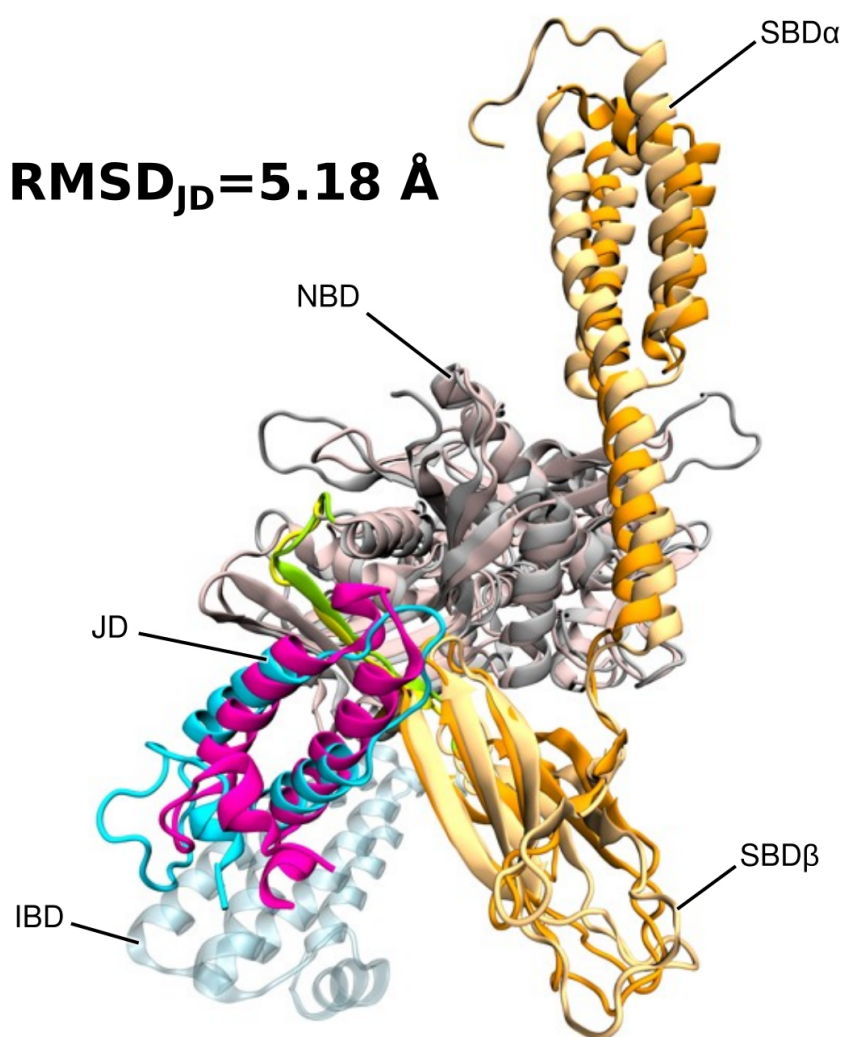
Zidentyfikowane struktury kompleksów Jac1-Ssq1



RYSUNEK 4.4: Reprezentatywne struktury kompleksu Jac1-Ssq1 dla największych klastrów oraz ich sumaryczny udział procentowy w trajektoriach uzyskanych dla przewidywanych struktur kompleksu otrzymanych w ramach dokowania molekularnego.

33 najlepsze struktury kompleksu Jac1-Ssq1 przy pomocy dokowania molekularnego (zgodnie z kryteriami opisanymi w podrozdziale 3.5.1). Dla każdej z tych struktur przeprowadziłem symulacje o długości ponad 500 ns każda, otrzymując trajektorie o sumarycznej długości 21 μ s.

By scharakteryzować sposoby wiązania Jac1 do Ssq1, przeprowadziłem klastrowanie dla całości trajektorii uzyskanych dla przewidywanych struktur kompleksu, w którym porównywałem ułożenie domeny J białka Jac1 względem NBD Ssq1. Wyniki klastrowania wskazywały na występowanie jednego dominującego klastra obejmującego łącznie ponad 56% otrzymanych trajektorii (Rys. 4.4), który znacząco przewyższał rozmiarem inne z 1080 możliwych stanów wiązania (największe zawierały do 6% trajektorii). Co więcej, choć w ciągu prowadzonych symulacji już od początku 8 trajektorii odpowiadało dominującemu stanowi wiązania, w czasie około 500 ns dodatkowych 9 innych trajektorii odnalazło główny stan wiązania. Co istotne, podobieństwo między sposobem wiązania kompleksu Jac1-Ssq1 przewidywanym w ramach dokowania a tym otrzymanym w wyniku symulacji spontanicznego wiązania Jac1 do Ssq1 jest bardzo duże (średnie odchylenie kwadratowe położenia (RMSD) dla ciężkich atomów domeny J <0.25 nm). Wynik ten dodatkowo wspiera hipotezę, że prezentowana struktura kompleksu Jac1-Ssq1, zidentyfikowana przy pomocy dwóch niezależnych metod, odpowiada specyficznemu oddziaływaniu tych białek.

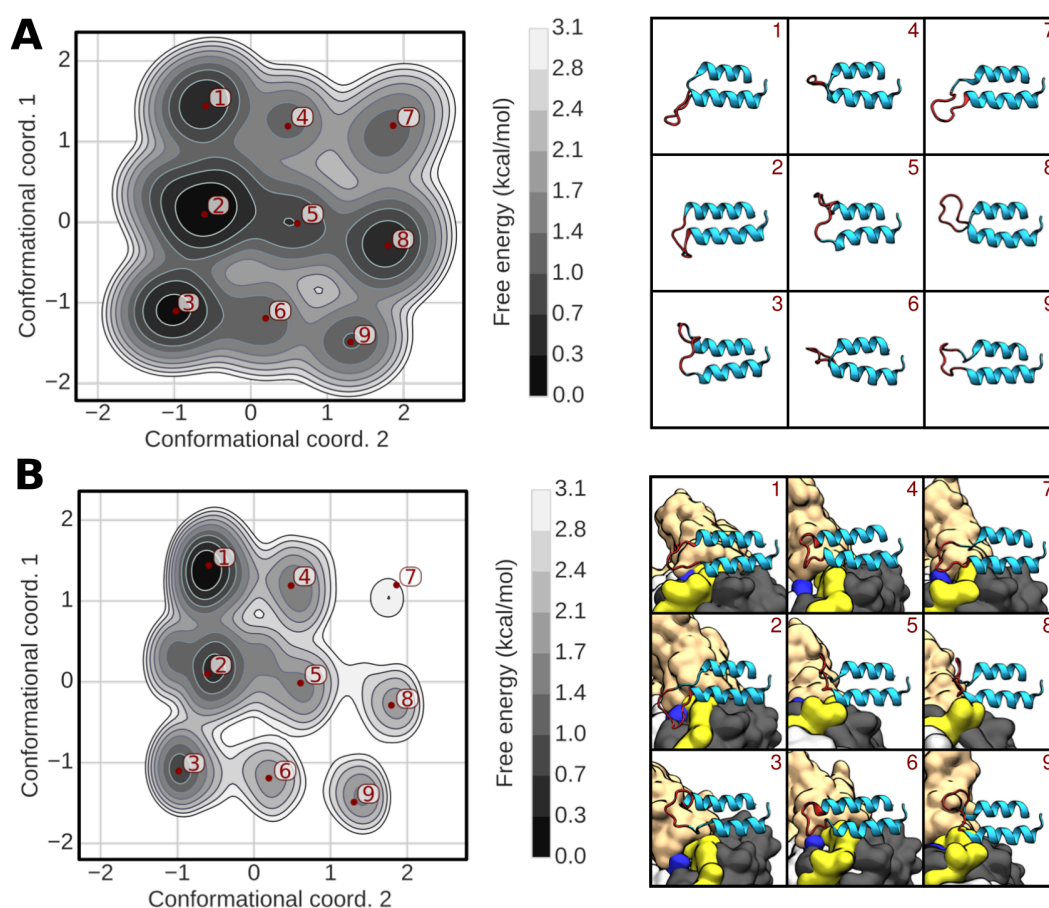


RYSUNEK 4.5: Nałożenie struktur kompleksów Jac1-Ssq1 oraz DnaJ-DnaK wraz z wartością RMSD wyznaczoną między domenami J (JD) względem położenia NBD białek HSP70.

W celu dodatkowej weryfikacji powyższych wyników, porównałem ułożenie domeny J względem NBD w dla stanu związanego Ssq1-Jac1 ze strukturą dla kompleksu białek DnaK-DnaJ wyznaczoną przy pomocy rentgenografii strukturalnej. [95] Jak widać na Rys. 4.5, obie domeny J białek JDP pozycjonują się w bardzo podobny sposób na powierzchni β -kartek subdomeny IIA NBD, co znajduje odzwierciedlenie w wartości RMSD 0.52 nm pomiędzy łańcuchami głównymi helis 2 i 3 domen J. Struktury kompleksów różnią się głównie w obszarze domeny J zawierającej motyw HPD, która w przypadku DnaJ jest w znaczącej mierze α -helikalna, zaś dla Jac1 przyjmuje strukturę nieuporządkowaną. W obu przypadkach jednak obie pętle sięgają powierzchni kontaktu domen NBD i SBD, znajdującego się w pobliżu łącznika między domenami.

4.1.3 Dopasowanie pętli HPD Jac1 do Ssq1 zachodzi zgodnie z mechanizmem selekcji konformacyjnej

Choć struktura białka Jac1 została wyznaczona metodą rengenografii strukturalnej, [179] to nie pozwoliła ona na zidentyfikowanie konformacji przyjmowanej przez pętlę HPD domeny J. Ponieważ pętla HPD odgrywa istotną rolę w stymulacji cyklu katalicznego Ssq1 przez Jac1, [120] postanowiłem scharakteryzować dynamikę konformacyjną tego fragmentu białka. W tym celu wykonałem zestaw 21 symulacji metadynamiki w wariacie well-tempered z wymianą replik (ang. bias-exchange metadynamics) o długości 400 ns każda, gdzie w 20 symulacjach potencjał przykładłem do jednego z kątów torsyjnych ϕ bądź ψ fragmentu pętli od reszty 53 do 63, zaś w ostatniej replice nie był przykładany żaden dodatkowy potencjał. Dzięki temu otrzymałem zbiór równowagowych konformacji pętli HPD białka Jac1, na podstawie których zidentyfikowałem możliwe stany konformacyjne pętli przy pomocy metody głównych składowych (PCA). W ten sposób udało mi się scharakteryzować 9 różnych stanów konformacyjnych pętli HPD dla niezwiązanego białka Jac1 (panel A Rys. 4.6). Stany te rozdzielone są przez niewielkie, wynoszące około 2 kcal mol^{-1}

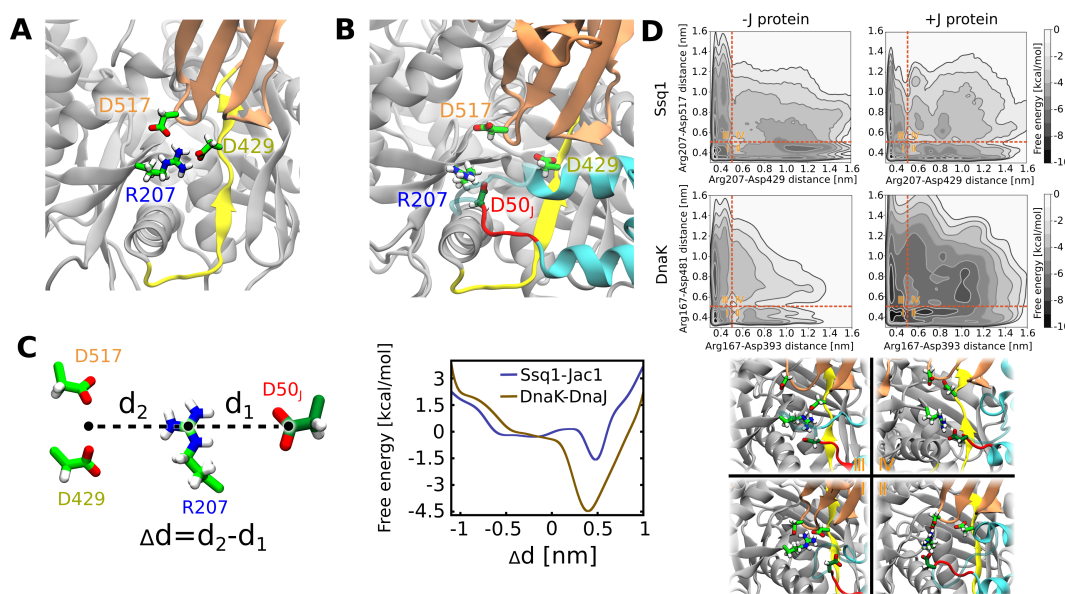


RYSUNEK 4.6: Mapy energii swobodnej zrzutowane na płaszczyznę największej zmienności konformacyjnej domeny J (lewa strona) oraz struktury odpowiadające minimom energii swobodnej (prawa strona) dla (A) wolnego białka Jac1 (B) Jac1 w kompleksie z Ssq1.

bariery energii swobodnej, umożliwiające szybką konwersję między nimi. W ramach tych stanów pętla HPD zarówno wyginała się powyżej, jak i poniżej płaszczyzny tworzonej przez helisy domeny J oraz skręcała się ku każdej z helis domeny J.

W kolejnym kroku sprawdziłem, w jaki sposób związanie do Ssq1 zmienia dynamikę konformacyjną pętli HPD Jac1. W tym celu zrzutowałem struktury domeny J otrzymane dla kompleksu Jac1-Ssq1 na wektory własne wyznaczone przy pomocy PCA dla wolnego białka Jac1. Jak można zauważyć na panelu B Rys. 4.6, pod wpływem związania pętla HPD głównie przyjmuje konformacje 1 i 2 (ponad 70% wszystkich stanów), w których sięga ona w kierunku powierzchni kontaktu domen NBD i SBD. Stabilizacja niektórych z uprzednio istniejących stanów pod wpływem związania do partnera wskazuje na to, że pętla HPD Jac1 wiąże się do Ssq1 zgodnie z mechanizmem tzw. selekcji konformacyjnej.

4.1.4 Pętla HPD białka JDP osłabia oddziaływanie pomiędzy domenami NBD i SBD- β poprzez oderwanie kluczowej arginy



RYSUNEK 4.7: **A** Reprezentatywna struktura białka Ssq1 prezentująca położenie kluczowej reszty Arg na powierzchni kontaktu NBD/SBD oraz mostki solne tworzone z resztą Asp łącznika oraz SBD. **B** Struktura reprezentująca spontaniczne utworzenie mostka solnego przez kluczową Arg z Asp motywu HPD kosztem oddziaływań na powierzchni kontaktu NBD/SBD. **C** Profile energii swobodnej dla zmiany partnerów oddziaływania kluczowej Arg między Asp motywu HPD a Asp łącznika i SBD dla Ssq1 (niebieski) oraz DnaK (złoty). **D** Mapy energii swobodnej dla zrywania oddziaływań między kluczową Arg a Asp łącznika bądź Asp SBD dla białek Ssq1 i DnaK wolnych (lewa strona) bądź w kompleksie z białkami JDP (prawa strona) oraz reprezentatywne struktury wyróżnionych stanów na przykładzie kompleksu Ssq1-Jac1.

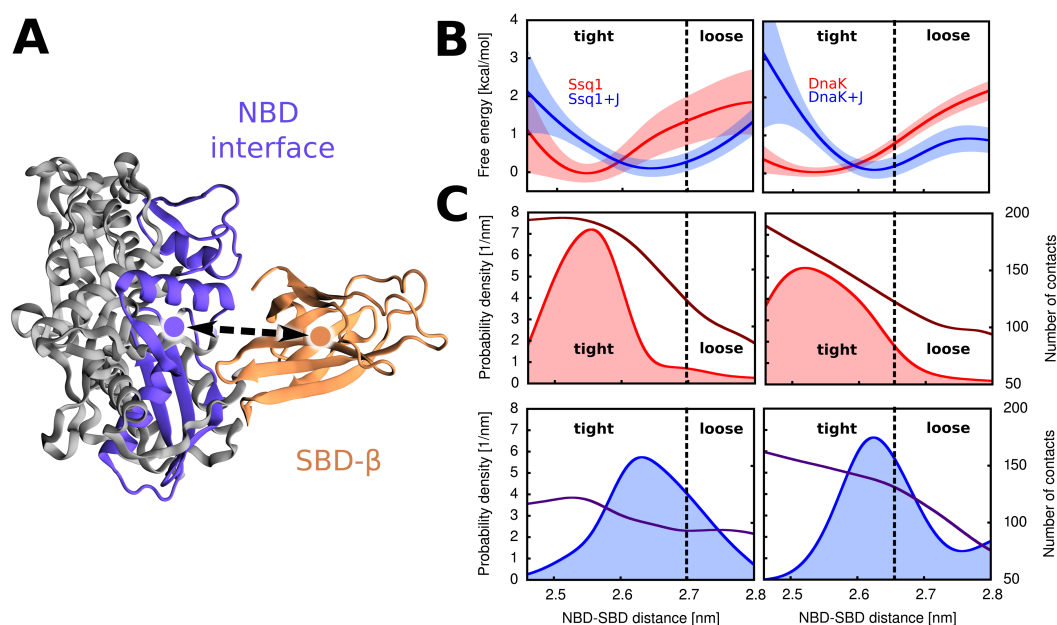
Preferowane konformacje pętli HPD, zidentyfikowane dla kompleksu Jac1 z Ssq1, lokalizują motyw HPD w pobliżu reszty Arg207 białka Ssq1, która jest silnie zakonserwowana ewolucyjnie wśród białek Hsp70 oraz jest konieczna dla prawidłowej komunikacji między NBD i SBD. [94, 113] Jak mogłem zaobserwować dla przeprowadzonych symulacji równowagowych białka Ssq1, Arg207 tworzy dwa mostki solne – jeden z Asp429 znajdującym się na łączniku domen, drugi z Asp517 zlokalizowanym na SBD- β (panel A Rys. 4.7). Choć analogiczny mostek solny między SBD- β a NBD (Arg167–Asp481) jest również obserwowany w strukturze krystalograficznej bakteryjnego białka DnaK, [113] to jednak nie występuje tam mostek solny Arg z łącznikiem. By zweryfikować tę rozbieżność, wyznaczyłem dwuwymiarowe mapy energii swobodnej zarówno dla wolnego białka Ssq1 jak i DnaK dla kontaktów między kluczową Arg (Arg207 dla Ssq1, Arg167 dla DnaK) a resztami na łączniku (Asp429 dla Ssq1, Asp393 dla DnaK) oraz na SBD- β (Asp517 dla Ssq1, Asp481 dla DnaK). W tym celu wykorzystałem symulacje metadynamiki w wariantcie well-tempered, biorąc jako współrzędne odległości między centrum masy grupy guanidynowej Arg207/167 a grupą karboksylową Asp429/393 bądź Asp517/481. W ramach tych symulacji skróciłem domenę SBD- α do fragmentu związanego z NBD (usuwając reszty 570–657 dla Ssq1 oraz 535–602 dla DnaK), znacząco zmniejszając rozmiar symulowanego układu. Dzięki temu mogłem przeprowadzić dłuższe symulacje pozwalające na uzbieźnienie uzyskanych map energii swobodnej. Otrzymane mapy zarówno dla izolowanego Ssq1 jak i DnaK (panel D Rys. 4.7, lewa kolumna) posiadają głębokie minimum dla niewielkich odległości między kluczową Arg a oboma Asp należącymi do łącznika i SBD, świadczące o tym, że frakcja obu białek z Arg tworzącą oba mostki solne zdecydowanie przeważa (82 % dla Ssq1 i 83 % dla DnaK) nad frakcjami białek z zerwanym przynajmniej jednym z mostków. Co zaskakujące, w przypadku DnaK frakcja białka z zerwanym mostkiem solnym między Arg a łącznikiem była bardzo mała $<1\%$. Bezpośrednie oddziaływanie między resztami Asp393 a Arg167 białka DnaK tłumaczyłoby zarówno dlaczego reszta Asp393 jest silnie zakonserwowana wśród białek Hsp70 [180] oraz dlaczego jest kluczowa dla komunikacji między domenami. [181]

W kolejnym kroku sprawdziłem, w jaki sposób związanie białek JDP do HSP70 osłabia poszczególne mostki solne tworzone przez kluczową Arg na powierzchni kontaktu NBD/SBD. W tym celu wyznaczyłem analogiczne mapy energii swobodnej kontaktów Arg–SBD- β i Arg–łącznik dla otrzymanej w tej pracy struktury kompleksu Ssq1 z Jac1 oraz dla struktury krystalicznej DnaK w kompleksie z domeną J białka DnaJ (panel D Rys. 4.7, prawa kolumna). [95] Pod wpływem związania białek JDP, prawdopodobieństwo utworzenia dwóch mostków solnych przez Arg znacząco spada w obu przypadkach (o 20 % w przypadku Ssq1 oraz o 32 % w przypadku DnaK). W przypadku Ssq1 przede wszystkim wzrastają frakcje z zerwanym jednym z mostków – odpowiednio 9 % i 11 % dla frakcji z zerwanym mostkiem Arg207–Asp429 oraz Arg207–Asp517. Z kolei w przypadku DnaK najmocniej osłabiane jest oddziaływanie między Arg167 a Asp393, dla którego frakcja z zerwanym mostkiem solnym między Arg a łącznikiem wzrasta z $<1\%$ do łącznie 35%. Sumarycznie, powyższe wyniki wskazują, że związanie białek JDP do Hsp70 osłabia oddziaływania na powierzchni kontaktu NBD/SBD.

Powtórnie analizując trajektorie dla głównego klastra kompleksu Jac1-Ssq1, zaobserwowałem spontaniczne oddysocjowanie Arg207 od swoich partnerów wiązania na rzecz utworzenia oddziaływań z Asp50 motywu HPD (panel B Rys. 4.7). Bazując na tej obserwacji uznałem, że białka JDP mogą stymulować postęp cyklu katalitycznego Hsp70 poprzez utworzenie mostka solnego między motywem HPD

a kluczową Arg, co działałoby się kosztem oddziaływań między NBD i SBD- β . By zeweryfikować tę hipotezę, wykonałem symulacje metadynamiki w wariantcie well-tempered, gdzie jako współrzędną opisującą wybrałem różnicę dwóch odległości: 1. pomiędzy centrami mas grup karboksylowych Asp łącznika i SBD- β (Asp 429 i 517 dla Ssq1, Asp393 i 481 dla DnaK) a centrum masy grupy guanidynowej kluczowej Arg oraz 2. pomiędzy centrum masy grupy karboksylowej Asp motywu HPD białek JDP (Asp50 dla Jac1, Asp35 dla DnaJ) a centrum masy grupy guanidynowej kluczowej Arg. Dodatkowo, w każdym z przypadków utrzymywałem odległości między wybranymi Asp a Arg poniżej 1.5 nm. Na obu wyznaczonych profilach energii swobodnej (panel C Rys. 4.7) można zaobserwować wyraźne, głębokie minimum (około $-1.5 \text{ kcal mol}^{-1}$ dla Ssq1 i ponad -4 kcal mol^{-1} dla DnaK) przy wartości współrzędnej 0.5 nm, świadczące o preferencji Arg do tworzenia mostka solnego z Asp motywu HPD względem tworzenia mostków solnych z resztami Asp na SBD- β i łączniku jednocześnie. Wynik ten wspiera naszą hipotezę, zgodnie z którą motyw HPD zrywa kontakty kluczowej Arg na powierzchni kontaktu NBD z SBD- β .

4.1.5 Związywanie domeny J do HSP70 powoduje wstępne rozdysocjowanie domen NBD i SBD- β



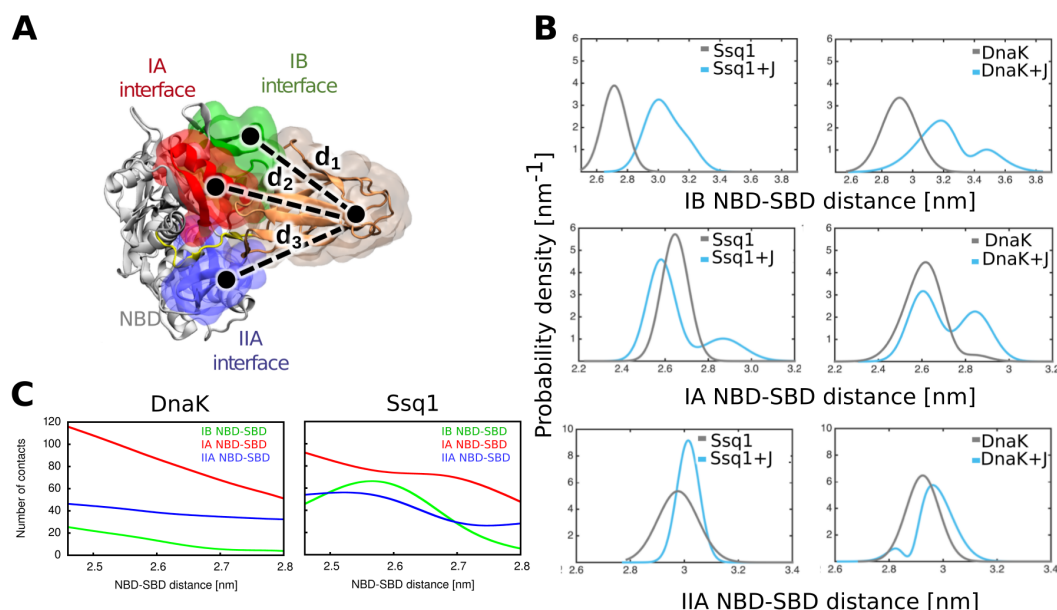
RYSUNEK 4.8: **A** Współrzędna reakcji dla rozsunięcia domen NBD i SBD- β . **B** Profile energii swobodnej dla Ssq1 (lewa strona) bądź DnaK (prawa strona) w stanie wolnym (czerwony) bądź w kompleksie z białkiem JDP (niebieski). **C** Rozkład prawdopodobieństwa dla odległości między domenami NBD i SBD dla białek Ssq1 i DnaK w stanie wolnym (czerwone) oraz w kompleksie z białkiem JDP (niebieskie). Ciemniejszym kolorem zaznaczono profil średniej liczby kontaktów między domenami wzdłuż odległości między domenami.

Skoro związywanie domeny J do HSP70 powodowało osłabienie sieci oddziaływań

tworzonych między domenami NBD i SBD- β , postanowiliśmy sprawdzić czy zwiążanie białek JDP wpływa na wzajemne ułożenie domen NBD i SBD względem siebie. W tym celu na początku wyznaczyłem profile energii swobodnej dla rozseparowania domen NBD i SBD- β białek Ssq1 i DnaK w stanie wolnym oraz w stanie zwiążanym z białkami JDP przy pomocy techniki umbrella sampling. [166] Jako współrzędną reakcji wybrałem odległość między środkami mas wyznaczonymi dla atomów węgla α domeny SBD- β oraz dla atomów węgla α fragmentów NBD znajdujących się w pobliżu SBD- β (reszty 112–143 oraz 186–262 dla Ssq1, reszty 75–104 oraz 146–226 dla DnaK). W ramach tych symulacji usunąłem całą domenę SBD- α , by zarówno odwzorować stan, w którym domena ta jest oddysocjowana od NBD, jak i dokładniej wypróbować zmianę kontaktów między NBD i SBD- β (dzięki możliwości przeprowadzenia dłuższych symulacji zmniejszonego układu). Jak przedstawiono na Rys. 4.8, profile energii swobodnej dla odległości między NBD i SBD- β zarówno dla DnaK i Ssq1 wykazują minimum dla małych odległości (2.55 nm dla Ssq1 i 2.52 nm dla DnaK), odpowiadające ściślemu zwiążaniu obu domen. Pod wpływem zwiążania białka JDP w obu przypadkach minimum przesuwają się ku większym wartościom o około 0.1 nm. Chcąc sprawdzić, w jakim stopniu obserwowane rozsuniecie domen odpowiada zmianie powierzchni kontaktu NBD i SBD- β , wyznaczyłem liczbę kontaktów tworzonych pomiędzy domenami (definiowaną jako odległość między resztami domen nie większą niż 0.5 nm) dla obu Hsp70 w stanie wolnym oraz w kompleksie z białkiem JDP (Rys. 4.8). Zwiążanie białka z domeną J zmniejsza średnią liczbę kontaktów między NBD a SBD- β o 41 % w przypadku Ssq1 (z 172 do 102) oraz o 21% w przypadku DnaK (z 154 do 122), co świadczy o znacznej reorganizacji położenia domen względem siebie. Co istotne, w stanie wolnym frakcje białka o tak niskiej liczbie kontaktów między domenami (odległości między domenami powyżej 2.7 nm dla Ssq1 oraz 2.66 nm dla DnaK) są bardzo małe – tylko około 5%. Sumarycznie, otrzymane wyniki wskazują, że utworzenie kompleksu z białkiem JDP indukuje przesunięcie ku konformacji Hsp70, w których SBD- β jest luźniej zwiążane z NBD.

By dokładniej scharakteryzować obserwowaną zmianę konformacyjną w obrębie powierzchni kontaktu NBD/SBD- β sprawdziłem, w jaki sposób zmienia się ułożenie SBD względem każdej z subdomen NBD wchodzących w skład współrzędnej reakcji (subdomeny IA, IB i IIA). W tym celu wyznaczyłem rozkłady prawdopodobieństwa dla odległości między SBD a każdą z subdomen na podstawie trajektorii otrzymanych w ramach symulacji US. Jak można zauważyć na panelu B Rys. 4.9, w obu przypadkach zwiążanie białka JDP do Hsp70 indukuje przede wszystkim rozsuniecie SBD z subdomeną IB oraz pojawienie się dodatkowego stanu z SBD odsuniętym od subdomeny IA zawierającej kluczową Arg, zaś odległość SBD od subdomeny IIB pozostaje podobna.

To, że zmiana powierzchni kontaktu wynika w przeważającej mierze z rozsunienia SBD i subdomeny IB jest dość zaskakujące, gdyż jest to subdomena tworząca najmniejszą liczbę kontaktów z SBD (panel C Rys. 4.9). Z tej przyczyny sprawdziłem, czy zwiążanie białek JDP do Hsp70 powoduje zmiany w obrębie samej domeny NBD. W tym celu wyznaczyłem rozkłady prawdopodobieństwa dla kąta torsyjnego między centrami mas subdomen IB-IA-IIA-IB dla Hsp70 w stanie wolnym oraz w kompleksie z białkiem JDP. Zwiążanie JDP do Hsp70 indukuje zmniejszenie się kąta torsyjnego między płacami I i II NBD o około 5° (Rys. 4.10). Sprawdziłem dalej, czy obserwowana zmiana ułożenia płaców NBD odpowiada częściowemu przejściu ze stanu NBD zwiążanego z ATP do stanu zwiążanego z ADP. W tym celu najpierw przeprowadziłem dodatkową 2 μ s symulację białka DnaK zwiążanego z

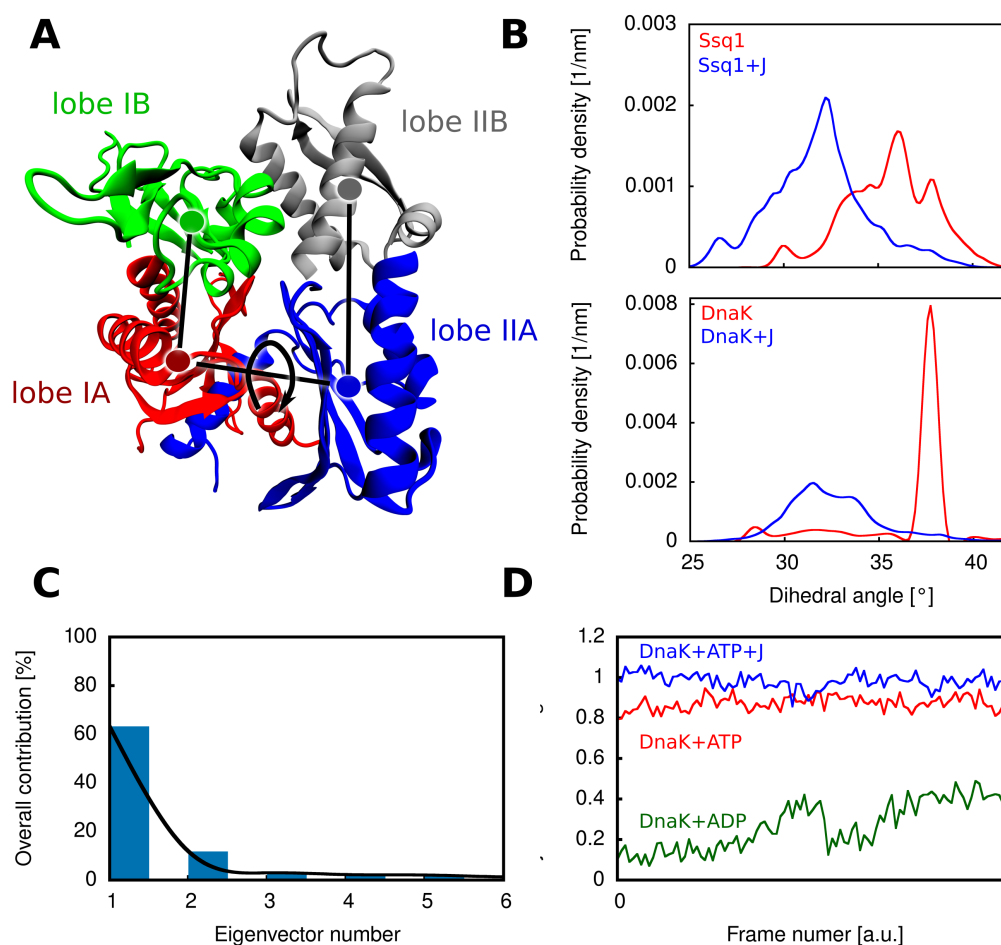


RYSUNEK 4.9: **A** Definicje odległości między subdomenami NBD a SBD- β . **B** Rozkłady prawdopodobieństwa odległości między subdomenami NBD a SBD- β dla wolnych białek Hsp70 oraz w kompleksie z białkami JDP. **C** Profile liczby kontaktów między subdomenami NBD a SBD- β dla wolnych białek Ssq1 i DnaK.

ADP o konformacji zamkniętej.[115] W kolejnym kroku wykorzystałem trajektorie uzyskane dla konformacji zamkniętej i otwartej białka DnaK do scharakteryzowania zmiany konformacyjnej NBD zachodzącej pod wpływem wymiany nukleotydu przy pomocy analizy głównych składowych. Jak można zauważyć na panelu C Rys. 4.10, pierwsza składowa opisuje ponad 64% zmienności układu, wystarczając do opisu przejścia konformacyjnego między stanami konformacyjnymi NBD. W kolejnym kroku zrzutowałem trajektorię otrzymaną dla kompleksu DnaK z domeną J białka DnaJ na wyznaczony wcześniej pierwszy wektor własny dla zmiany konformacyjnej NBD. Jak można zobaczyć na panelu D Rys. 4.10, związanie domeny J do DnaK nie indukuje przejścia konformacyjnego od struktury NBD związanej z ATP do struktury NBD z ADP, a raczej wspomaga jeszcze bardziej kompaktowe ułożenie NBD. Powyższy wynik sugeruje, że związanie białek JDP przez Hsp70 najprawdopodobniej ułatwia przyjęcie przez NBD konformacji kompetentnej do hydrolizy ATP.

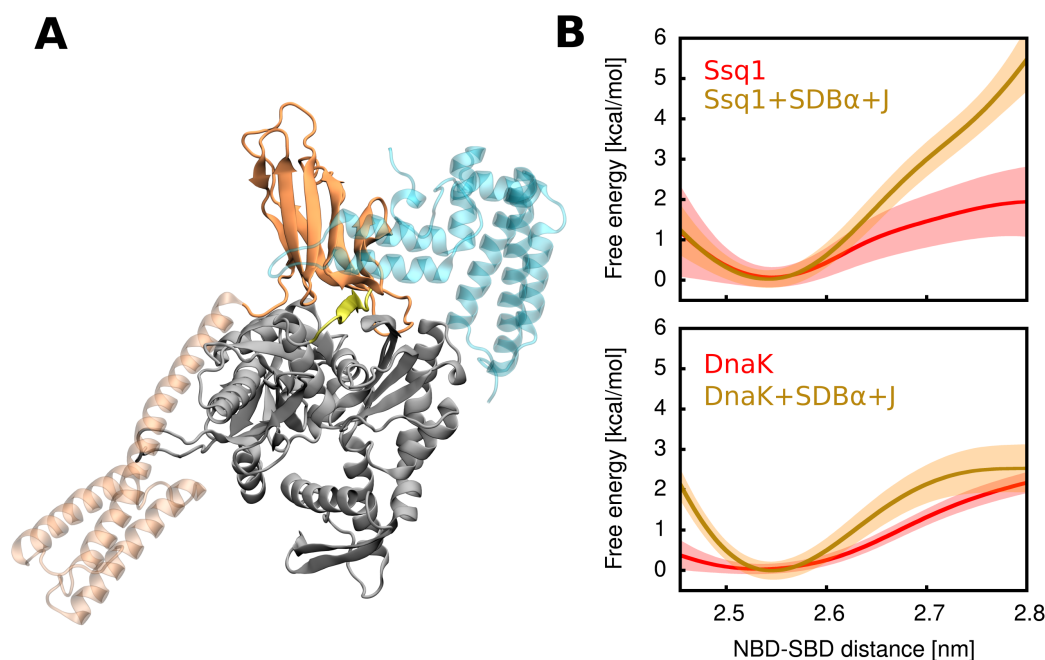
W kolejnym kroku sprawdziłem, czy oddysocjowanie SBD- α jest konieczne dla zaobserwowania zmiany powierzchni kontaktu NBD/SBD- β . W tym celu wyznaczyłem profile energii swobodnej względem odległości między domenami dla kompleksów białek JDP z pełnymi wariantami białek Hsp70. Zarówno w przypadku DnaK jak i Ssq1 nie jest obserwowana tendencja do poluzowania kontaktów między NBD/SBD- β , gdy SBD- α pozostaje zasocjowana z NBD (Rys. 4.11). Dodatkowo minimum energii swobodnej ulega zwężeniu, co sugeruje, iż zasocjowanie SBD- α do NBD ułatwia utrzymanie domen NBD i SBD- β w ściśle zdefiniowanej orientacji względem siebie. Powyższe wyniki sugerują również rolę substratu w stymulacji cyklu katalitycznego, który wiążąc SBD- α umożliwiałby zmiany w obrębie powierzchni kontaktu NBD/SBD- β .

Chcąc sprawdzić rolę obserwowanej zmiany powierzchni kontaktu NBD/SBD- β w



RYSUNEK 4.10: **A** Schematyczne przedstawienie kąta torsyjnego między subdomenami NBD. **B** Rozkłady prawdopodobieństwa kąta torsyjnego między subdomenami NBD dla wolnych białek Hsp70 (czerwony) oraz w kompleksie z białkami JDP (niebieski). **C** Wkłady kolejnych wektorów własnych do opisu całkowitej zmienności w obrębie łańcucha głównego NBD. **D** Przebiegi trajektorii symulacji równowagowych dla białka DnaK konformacji zamkniętej (zielony), otwartej (czerwony) oraz otwartej w kompleksie z domeną J białka DnaJ (niebieski) rzutowane na pierwszy wektor własny.

postępie cyklu katalitycznego białek Hsp70, sprawdziłem czy znane mutacje upośledzające komunikację między domenami NBD i SBD- β wpływają na ułożenie domen względem siebie. W tym celu wyznaczyłem profile energii swobodnej względem odległości między domenami dla wariantów Ssq1 i DnaK, w których kluczowa Arg została zastąpiona przez Asp. Ponadto analizowałem również wariant mutacyjny DnaK Arg167Cys Asp481Cys z utworzonym mostkiem disiarczkowym między tymi resztami, który uniemożliwia pełne rozdysocjowanie domen. Jak przedstawiono na Rys. 4.12, doliny energii swobodnej znacząco się poszerzają dla każdej z wybranych mutacji, pozwalając na przyjmowanie konformacji w których SBD- β jest luźno związane z NBD. Mutacje te powodują około 5-krotny wzrost frakcji białka o luźno związanych domenach w przypadku mutacji Arg do Asp oraz 2,5-krotny w przypadku mutantu z mostkami disiarczkowymi. Wyniki doświadczeń biochemicznych otrzymane w zespole prof. Marszałka wskazały, że mutacja Arg207Asp

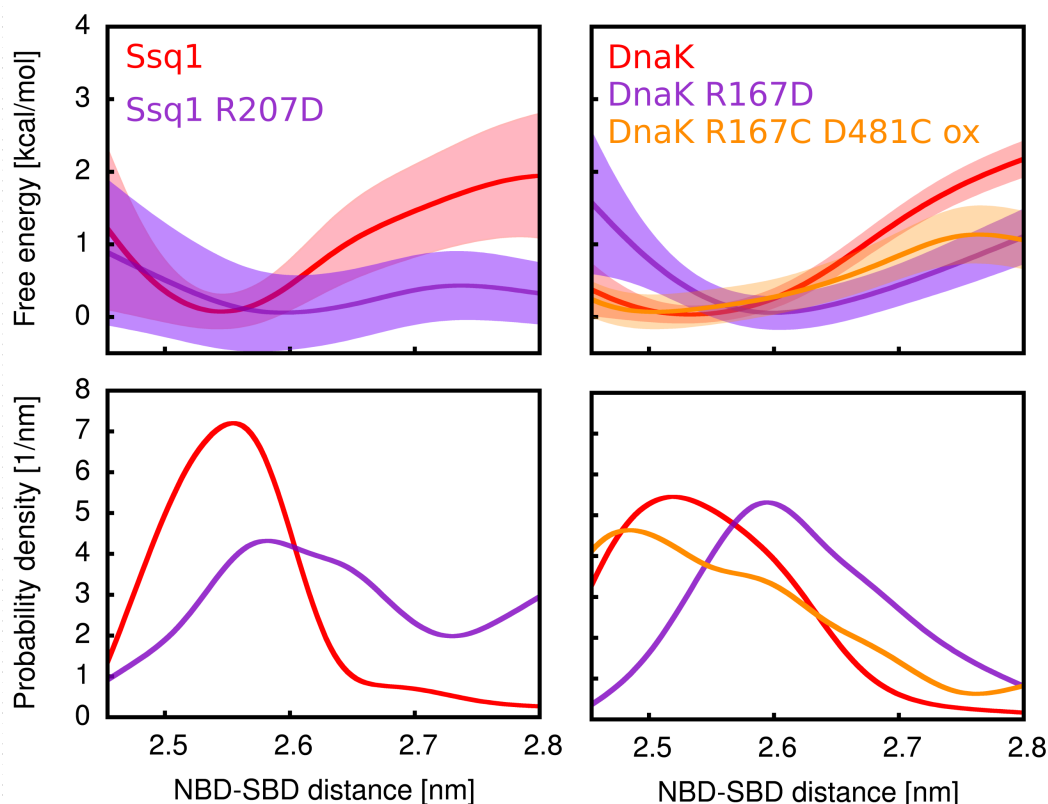


RYSUNEK 4.11: **A** Schematyczne przedstawienie porównywanych układów: pierwszy złożony jest z NBD i SBD- β (matowy), w drugim dodatkowo obecne są SBD- α i białko JDP (przezroczyste). **B** Profile energii swobodnej dla rozejścia się domen NBD i SBD- β Ssq1 (górze) bądź DnaK (dół) w stanie wolnym bez SBD- α (czerwony) bądź z SBD- α w kompleksie z białkiem JDP (żółty).

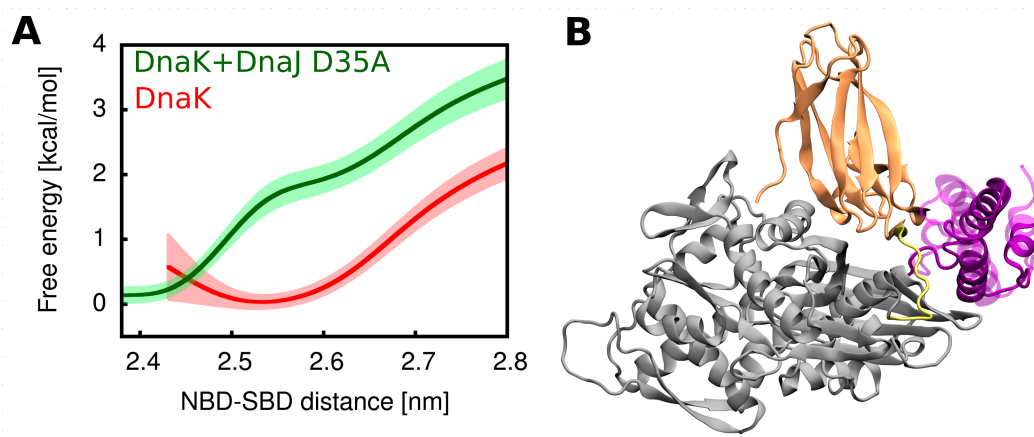
białka Ssq1 skutkuje ponad trzykrotnym wzrostem bazowej aktywności ATPazy, jak również praktycznie całkowitym zniesieniem stymulacji ATPazy przez białko Jac1. Powyższe wyniki wspierają hipotezę, zgodnie z którą poluzowanie kontaktów między NBD i SBD- β umożliwia postęp cyklu katalitycznego białek Hsp70.

4.1.6 Steryczne odpychanie między helisą 3 domeny J a SBD skutkuje wstępnym rozdysocjowaniem domen SBD i NBD białek Hsp70

Motyw HPD odgrywa kluczową rolę w stymulacji ATPazy białek Hsp70 przez białka JDP.[120, 182] Postanowiłem zatem sprawdzić na przykładzie kompleksu DnaK z domeną J białka DnaJ, w jaki sposób usunięcie specyficznego oddziaływania między motywem HPD a kluczową argininą, wpłynie na indukowanie zmiany konformacyjnej obserwowanej dla Hsp70. W tym celu wyznaczyłem profil energii swobodnej dla rozejścia domen NBD i SBD dla białka DnaK w kompleksie z mutantem Asp35Ala domeny J białka DnaJ. Jak można zaobserwować na panelu A Rys. 4.13, mutant Asp35Ala domeny J nie indukuje zmiany konformacyjnej Hsp70, skutkującej luźniejszym w obrębie domen NBD i SBD białka DnaK. Dalsza analiza trajektorii uzyskanych w ramach symulacji US wskazała na odmienne ułożenie domeny J na powierzchni NBD w przypadku mutantu Asp35Ala w porównaniu z białkiem dzikiego typu DnaJ: w przypadku mutantu w ścisłym kontakcie z DnaK pozostawała wyłącznie helisa 2 domeny J, zaś helisa 3 wyraźnie oddalała się od SBD- β (panel B Rys. 4.13).

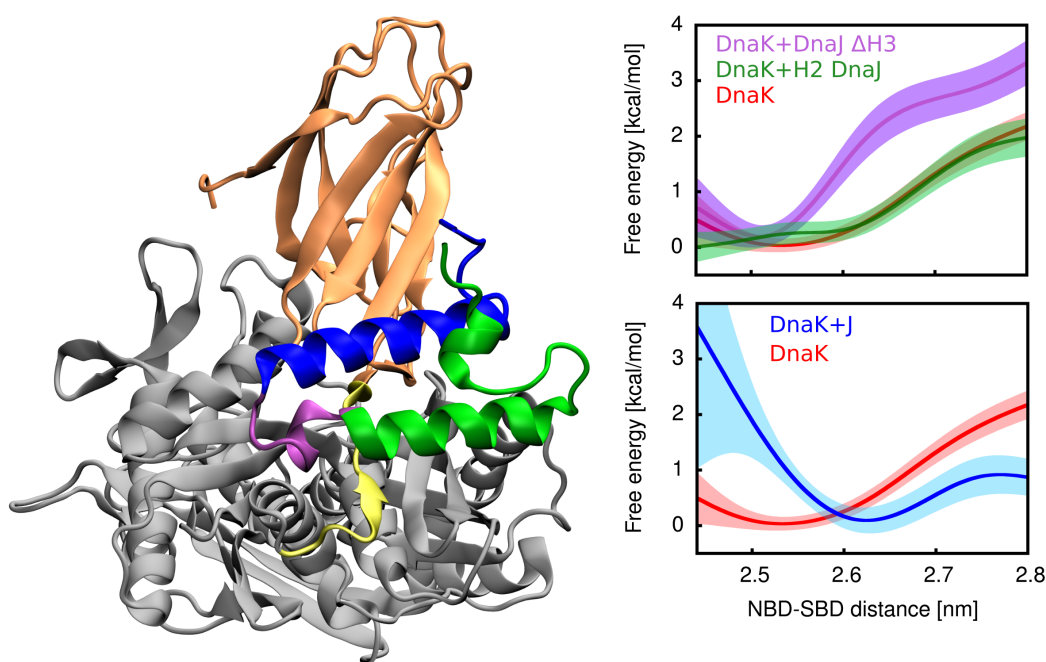


RYSUNEK 4.12: (góra) Profile energii swobodnej dla rozejścia domen NBD i SBD- β dla białek Ssq1 (lewa strona) bądź DnaK (prawa strona) oraz wybranych mutantów. (dół) Rozkłady gęstości prawdopodobieństwa względem odległości między NBD i SBD- β odpowiadające zaprezentowanym powyżej profilom energii swobodnej.



RYSUNEK 4.13: **A** Profile energii swobodnej dla rozejścia domen NBD i SBD- β dla białka DnaK w stanie wolnym (czerwone) bądź w kompleksie z mutantem Asp35Ala białka J (zielony). **B** Struktury kompleksu DnaK z mutantem Asp35Ala domeny J białka DnaJ (matowy) bądź szcpepem dzikim domeny J (przezroczysty).

Powyższa obserwacja stanowiła punkt wyjścia dla postawienia następującej hipotezy: o ile helisa 2 odpowiadałaby za rozpoznanie miejsca wiązania na obszarze NBD, oddziaływanie między motywem HPD a kluczową Arg zarówno osłabiałoby



RYSUNEK 4.14: **A** Schematyczne przedstawienie podziału domeny J białka DnaJ: fragment zielony oznaczany jest jako H2, zielony i magenta jako $\Delta H3$, cała domena J jako +J. **B** Profile energii swobodnej dla rozejścia domen NBD i SBD β dla DnaK w stanie wolnym (czerwony), w obecności H2 (zielony) oraz $\Delta H3$ (fioletowy). Dla porównania poniżej załączono profile dla rozdysocjowania domen DnaK w stanie wolnym (czerwony) oraz związanego z całą domeną J białka DnaJ (niebieski).

oddziaływania na powierzchni kontaktu NBD/SBD- β , jak również pozycjonowałyby domenę J w ściśle zdefiniowany sposób względem NBD i SBD- β , zaś helisa 3 odpowiadałaby wyłącznie za steryczne odpychanie SBD- β , wymuszając zarówno zmianę obserwowaną w obrębie NBD, jak i rozsuniecie NBD i SBD- β . By zweryfikować tak sformułowaną hipotezę, postanowiłem wyznaczyć profile energii swobodnej względem odległości między domenami dla białka DnaK w kompleksie kolejno z domeną J białka DnaJ, domeną J pozbawioną helisy 3 oraz samą helisą 2 domeny J. Porównanie tych profili pozwoliłoby na określenie, jaki jest wystarczający fragment domeny J do indukowania zmiany powierzchni kontaktu NBD/SBD- β . By umożliwić miarodajne porównanie między systemami, zarówno domenę J pozbawioną helisy 3 oraz samą helisę 2 domeny J utrzymywałem w konformacji odpowiadającej pełnej domenie J przy pomocy dodatkowego potencjału. Jak przedstawiono na Rys. 4.14, ani związanie helisy 2 domeny J, ani związanie domeny J pozbawionej helisy 3 do Hsp70 nie skutkuje przyjmowaniem konformacji o luźniej związanych domenach NBD i SBD- β . Kształt profilu energii swobodnej dla rozdysocjowania domen w przypadku obu kompleksów DnaK z fragmentami domeny J jest porównywalny z profilem uzyskanym dla wolnego białka DnaK. Powyższe wyniki sugerują, że o ile hydrofilowa helisa 3 domeny J nie tworzy specyficznych kontaktów z SBD- β , to jej obecność jest konieczna do sterycznego odpychania SBD- β od NBD.

4.2 Zmiany strukturalne domeny 4HB przysposabiające do aktywacji genów indukujących oporność wielolekową u drożdży *Saccharomyces cerevisiae*

4.2.1 W procesie ewolucji *Saccharomyces cerevisiae* znacząco wzrasta hydrofobowość C-terminalnego końca domeny 4HB

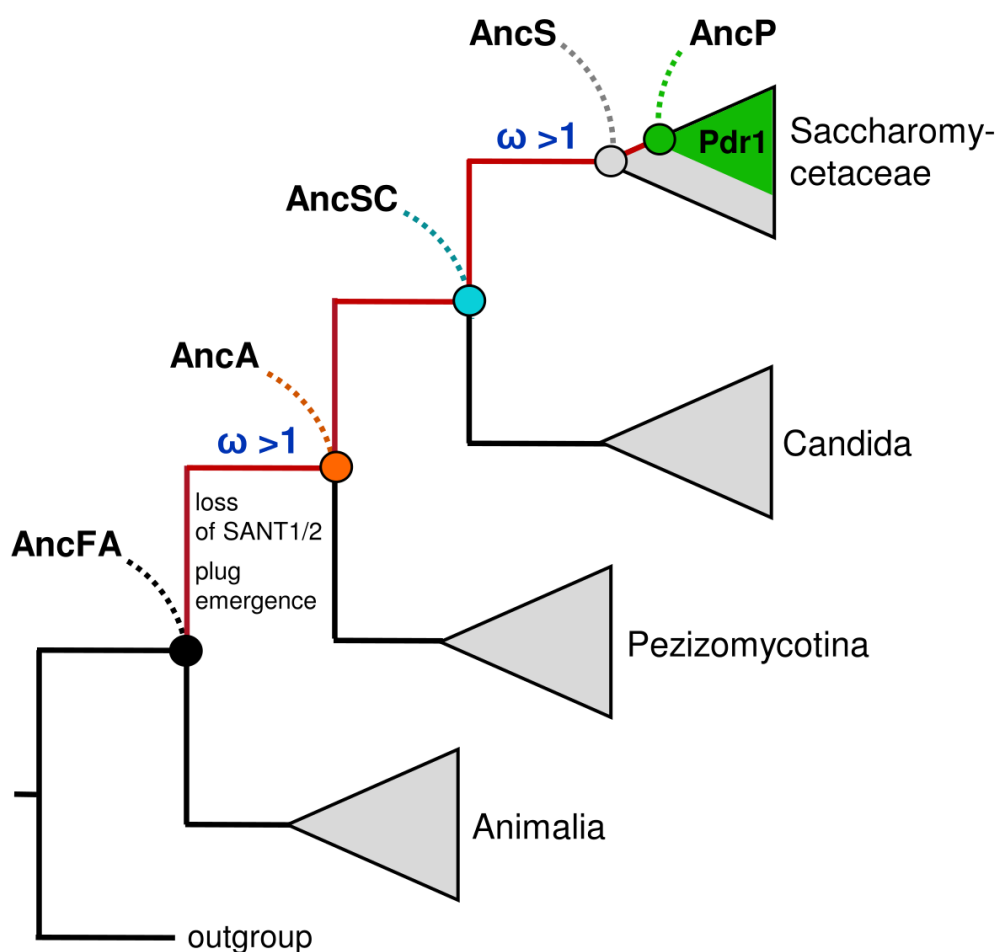
Badania zespołu prof. Craig wykazały, iż 13-resztowy C-terminalny fragment domeny 4HB wystarcza do aktywacji oporności wielolekowej u drożdży *S. cerevisiae*. [98] Fragment ten u drożdży przyjmuje strukturę nieuporządkowanego łańcucha, który dokuje do domeny 4HB między helisami 2 i 4. [98, 125] Analogiczny C-terminalny fragment nie występuje w innych organizmów, które na C-końcu zotytny posiadają domeny SANT. [131] U *S. cerevisiae* C-terminalny koniec jest bogaty w reszty hydrofobowe (aż 7 z 13). Każda z tych reszt jest konieczna dla aktywacji czynnika transkrypcyjnego odpowiedzialnego za ekspresję genów indukujących oporność wielolekową. [98] By sprawdzić, czy wysoka hydrofobowość C-terminalnego końca jest specyficzna dla drożdży wykazujących oporność wielolekową, przeprowadziłem symulacje o długości 1 μ s znanych struktur domen 4HB, na podstawie których wyznaczyłem wkłady poszczególnych helis oraz C-terminalnego końca do rdzenia hydrofobowego domeny 4HB. Wkład fragmentów do rdzenia hydrofobowego zdefiniowałem jako różnicę między całkowitą powierzchnią reszt hydrofobowych danego fragmentu a ich powierzchnią wyeksponowaną do wody. Jak pokazano na Tab. 4.1, zarówno w przypadku ludzkiej domeny 4HB, jak i grzybowej z *Chaetomium thermophilum*, reszty hydrofobowe są w miarę równomiernie rozłożone pomiędzy każdy z fragmentów białka, zaś hydrofobowość C-terminalnego końca jest nieznaczna. Przeciwnie jest w przypadku *S. cerevisiae*, gdzie główny wkład w hydrofobowość rdzenia domeny 4HB pochodzi od C-terminalnego fragmentu domeny. Sugeruje to, że wysoka hydrofobowość C-końca jest elementem przystosowawczym domeny 4HB do pełnienia dodatkowej funkcji aktywowania oporności wielolekowej.

TABLICA 4.1: Wkłady do powierzchni tworzącej rdzeń hydrofobowy znanych struktur domeny 4HB pochodzące od jej wybranych fragmentów.

	Powierzchnia hydrofobowa [nm ²]		
	<i>C. thermophilum</i>	<i>S. cerevisiae</i>	human
Helisa 1	2.03 ± 0.17	0.22 ± 0.05	1.22 ± 0.25
Helisa 2	0.61 ± 0.11	1.13 ± 0.28	1.03 ± 0.23
Helisa 3	0.52 ± 0.09	0.34 ± 0.09	0.50 ± 0.11
Helisa 4	1.32 ± 0.13	0.39 ± 0.09	0.52 ± 0.11
C-term	0.50 ± 0.11	2.19 ± 0.34	0.10 ± 0.07

By dodatkowo wesprzeć powyższy wniosek, postanowiliśmy sprawdzić, na którym etapie procesu ewolucji pojawił się wysoce hydrofobowy C-terminalny koniec.





RYSUNEK 4.15: Filogeneza domeny 4HB u *Saccharomycetaceae*, *Candida*, *Pezizomycotina*, *Animalia* oraz pierwotniaków (oznaczonych jako outgroup). Wspólni przodkowie oznaczeni są przez kropki: czarna - zwierząt i grzybów (AncFA), pomarańczowa - *Ascomycota* (AncA), błękitna - *Saccharomycetaceae* oraz *Candida* (AncSC), szara - *Saccharomycetaceae* (AncS), zielona *Saccharomycetaceae* wykazującego oporność wielolekową (AncP). $\omega > 1$ oznacza zmiany zachodzące pod wpływem pozytywnej presji ewolucyjnej.

W tym celu wyznaczyłem wkłady do hydrofobowego rdzenia dla przodków *S. cerevisiae* przewidzianych na podstawie analiz bioinformatycznych (przeprowadzonych przez dr. Tomiczka oraz przedstawionych za jego zgodą na Rys 4.15; przygotował on również modele homologiczne domen 4HB pochodzących z tych białek). Dla każdego z przodków przeprowadziłem symulacje o długości 1.8 μ s, z czego pierwsze 800 ns potraktowałem jako relaksację symulowanego układu. Jak widać na Tab. 4.2, wspólny przodek *Ascomycota* (AncA) po utracie domeny SANT posiada C-terminalny fragment 4HB o stosunkowo niskim wkładzie do hydrofobowości rdzenia domeny; znajduje się na nim również niewiele reszt hydrofobowych. Wywodzący się od AncA wspólny przodek *Saccharomycetaceae* oraz *Candida* (AncSC) posiada już istotny wkład do hydrofobowego rdzenia pochodzący od C-terminalnego końca, który później zostaje zachowany aż do organizmów wykazujących oporność wielolekową. Wysoka hydrofobowość C-terminalnego fragmentu pozostała wspólna zarówno dla drożdżaków *Saccharomycetaceae* i *Candida*. Kolejne

zmiany, zachodzące pod wpływem pozytywnej presji ewolucyjnej u przodka *Saccharomycetaceae* (AncS), a następnie przodka *Saccharomycetaceae* wykazującego oporność wielolekową (AncP), wiązały się ze spadkiem udziału reszt hydrofobowych helis 1–4 do hydrofobowego rdzenia białka.

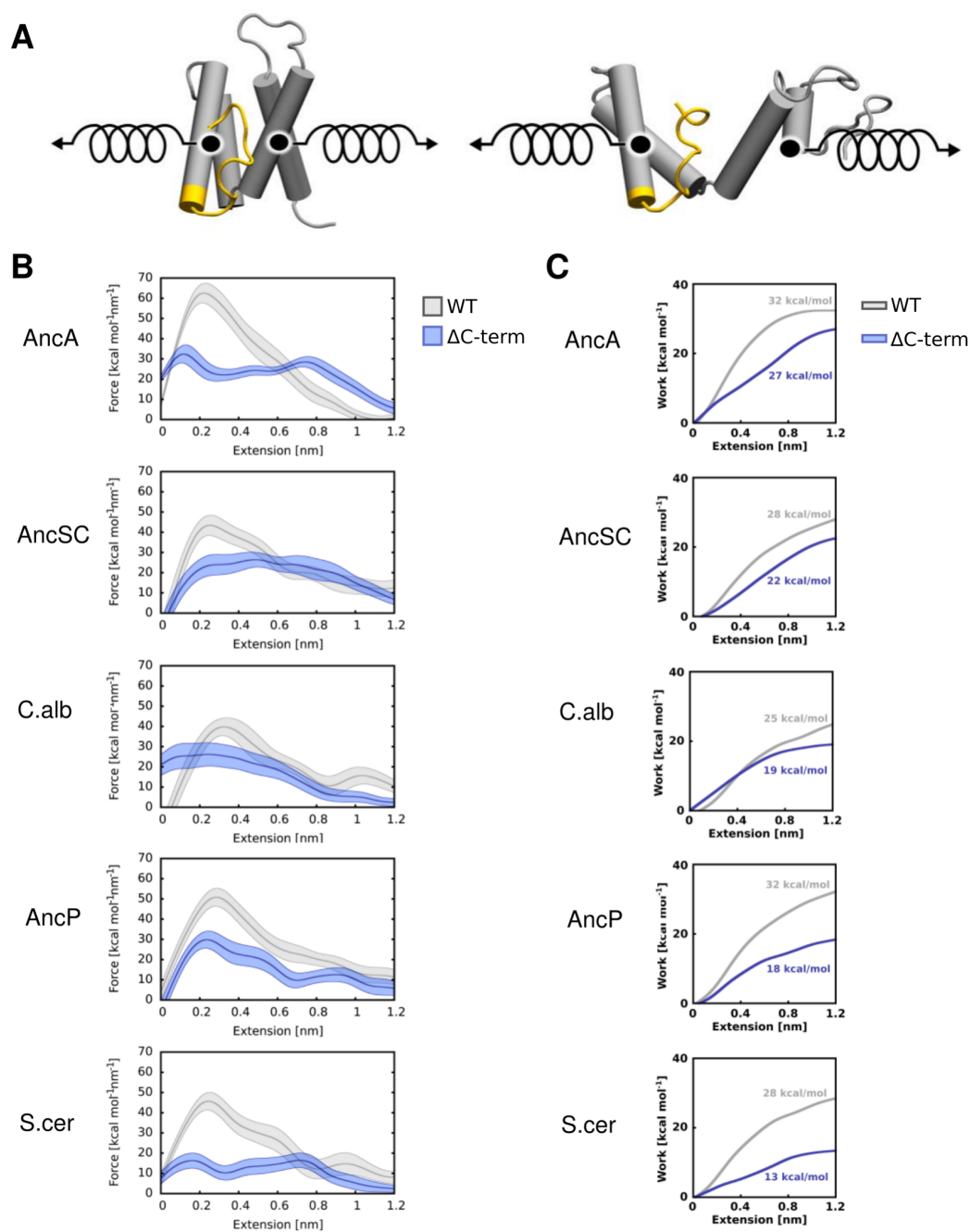
TABLICA 4.2: Wkłady do powierzchni tworzącej rdzeń hydrofobowy dla struktur domeny 4HB na linii filogenetycznej *S. cerevisiae* pochodzące od jej wybranych fragmentów.

	Powierzchnia hydrofobowa [nm ²]				
	<i>S. cerevisiae</i>	AncA	AncSC	<i>C. albicans</i>	AncP
Helisa 1	0.22 ± 0.05	2.02 ± 0.17	1.51 ± 0.35	1.52 ± 0.28	1.22 ± 0.21
Helisa 2	1.13 ± 0.28	0.69 ± 0.11	0.53 ± 0.10	1.60 ± 0.18	0.85 ± 0.43
Helisa 3	0.34 ± 0.09	0.38 ± 0.07	0.26 ± 0.07	0.25 ± 0.07	0.40 ± 0.10
Helisa 4	0.39 ± 0.09	1.65 ± 0.13	1.48 ± 0.13	1.51 ± 0.14	0.37 ± 0.08
C-term	2.19 ± 0.34	0.49 ± 0.08	2.37 ± 0.21	2.37 ± 0.25	1.72 ± 0.44

4.2.2 W procesie ewolucji związanie C-terminalnego końca do reszty domeny 4HB zaczęło odpowiadać za stabilizację tej domeny

Następujący wraz ze zmianami ewolucyjnymi wzrost wkładu C-terminalnego fragmentu 4HB do rdzenia hydrofobowego pozwala podejrzewać, że fragment ten zaczął w coraz bardziej istotny sposób kontrybuować do stabilizacji stanu zwiniętego domeny. Gdyby rzeczywiście tak było, to dysocjacja C-terminalnego fragmentu 4HB od helis skutkowałoby ułatwionym rozwinięciem domeny, co prowadziłoby do dodatkowej stabilizacji stanu z wyplecionym C-końcem. By sprawdzić, w jakim stopniu wkład C-terminalnego fragmentu 4HB zmieniał się w procesie ewolucji całej domeny, obliczyłem pracę konieczną do rozwinięcia domeny 4HB w wariacie pełnym (WT) oraz pozbawionym 11 reszt na C-końcu (Δ C-term). W tym celu dla każdej z domen białek ancestralnych (AncA, AncSC, AncP) oraz dla domen *S. cerevisiae* oraz *Candida albicans* (model homologiczny) przeprowadziłem po 6 symulacji dynamiki sterowanej, w trakcie których wymuszałem rozwijanie domen przy pomocy dodatkowego potencjału. Jako współrzędną opisującą rozwijanie domeny wybrałem odległość między centrami mas łańcuchów głównych helis 1 i 3 oraz helis 2 i 4.

W każdym przypadku początkowo siła wymagana do rozciągnięcia domen wzrasta (Rys. 4.16), co związane jest ze zrywaniem specyficznych oddziaływań odpowiadających stanowi zwiniętemu domeny. Po osiągnięciu maksimum oddziaływania te zostają zerwane, a siła konieczna do odciągnięcia fragmentów od siebie maleje wraz z postępującym rozwijaniem domeny. W każdym z analizowanych przypadków obecność fragmentu C-końcowego domeny przyczynia się do wzrostu maksimum siły wymaganej do rozciągnięcia domen. By w bardziej ilościowy sposób porównać proces rozwijania domen, scałkowałem profile sił, otrzymując wartość pracy wykonanej przez zewnętrzny potencjał w ramach rozwijania domen. Jak również pokazano na Rys. 4.16, zarówno w przypadku domen 4HB AncA, AncSC oraz *C. albicans*, fragment C-końcowy ma stosunkowo niewielki wkład (ok. 20%) do stabilności całej domeny 4HB, a część helikalna domeny pozostaje stabilna po usunięciu reszt C-końcowych. Sytuacja diametralnie zmienia się w przypadku AncP, gdzie usunięcie fragmentu C-końcowego obniża pracę wymaganą do rozwinięcia domeny o prawie



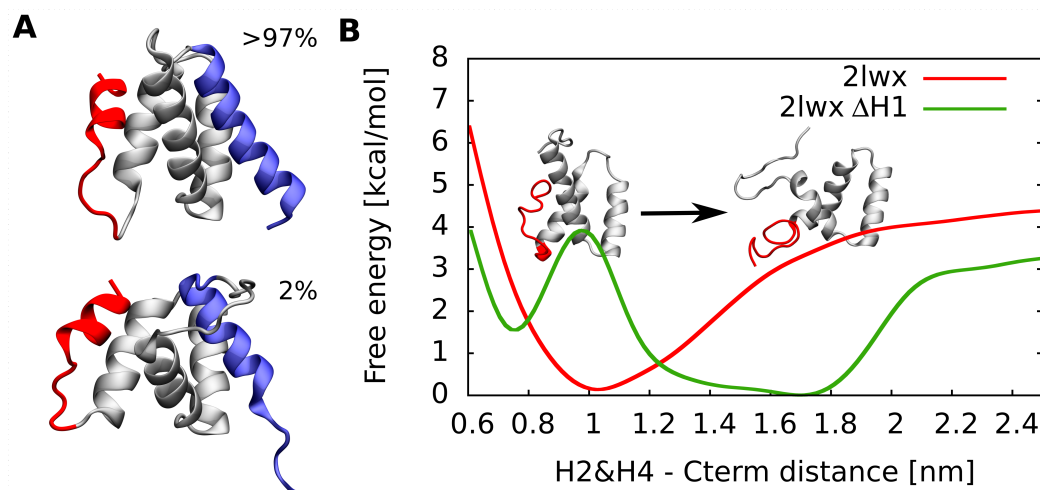
RYSUNEK 4.16: **A** Schematyczne przedstawienie współrzędnej opisującej proces rozwijania domeny 4HB. **B** Profile przykładowej siły względem rozciągnięcia oraz **C** pracy koniecznej do rozciągnięcia domeny dla pełnego wariantu domeny 4HB (WT, szare) bądź pozbawionego 11 terminalnych reszt (Δ C-term, niebieskie).

50%. W przypadku *S. cerevisiae* domena C-końcowa odpowiada już w zdecydowanej większości za stabilizację stanu zwiniętego domeny 4HB, zaś stabilność domeny helikalnej ulega dalszemu zmniejszeniu, o czym świadczą zarówno najmniejsza praca konieczna do rozwinięcia domeny dla wariantu Δ C-term, jak i dość płaski przebieg krzywej przyłożonej siły od rozciągnięcia domeny. Powyższe wyniki sugerują, że stabilność samodzielnego fragmentu Δ C-term domeny 4HB znacząco spada

w ramach ewolucyjnego przystosowania do pełnienia nowej funkcji.

4.2.3 Zmiany konformacyjne zachodzące w obrębie domeny 4HB są związane z wypętlaniem C-terminalnego fragmentu domeny na zewnątrz

Zgodnie z pomiarami przy pomocy technik NMR i CD dla termicznego rozwijania domeny 4HB *S. cerevisiae*, stabilność formy zwiniętej 4HB względem stanu rozwiniętego odpowiedzialnego za pełnienie dodatkowej funkcji jest dość niewielka – wynosi około -2 kcal mol^{-1} . [98] Postanowiłem zatem sprawdzić, czy symulacje molekularne pozwolą odtworzyć obserwowaną eksperymentalnie równowagę konformacyjną, a dzięki temu scharakteryzować konformacje 4HB odpowiadające za oddziaływanie z czynnikiem transkrypcyjnym aktywującym ekspresję genów odpowiedzialnych za oporność wielolekową. W tym celu wykonałem symulacje domeny 4HB metodą wymiany replik temperaturowych w wariacie (ang.) *solute tempering*, [154] stosując 11 replik w zakresie temperatur od 310 do 466 K, każda o długości 1.75 μs . Klastrowanie trajektorii uzyskanej w najniższej (fizjologicznej) temperaturze pozwoliło na zidentyfikowanie dwóch głównych stanów konformacyjnych domeny 4HB (Rys. 4.17). Pierwszy, stanowiący ponad 97% zespołu konformacyjnego, odpowiada strukturze w pełni zwiniętej białka. Drugi ze zidentyfikowanych stanów, stanowiący około 2% całości, różni się przede wszystkim w obszarze helisy 1, która ulega częściowemu rozwinięciu, oraz w obszarze C-końcowym, który jest bardziej wyeksponowany do środowiska wodnego. Wyniki te pozostają w dobrej zgodności z obserwacjami eksperymentalnymi, które wskazują zarówno na podobne proporcje między stanem zwiniętym a rozwiniętym (białko występuje w ok. 4% w stanie rozwiniętym w roztworze), jak również że oddysocjowanie helisy 1 jest konieczne dla pełnienia dodatkowej funkcji przez C-koniec domeny 4HB.[98]



RYSUNEK 4.17: **A** Reprezentatywne struktury domeny 4HB *S. cerevisiae* dla dwóch głównych klastrów oraz udział procentowy klastrów w trajektorii uzyskanej przy pomocy symulacji replik temperaturowych. **B** Profile energii swobodnej dla wyplecenia fragmentu 6 reszt C-końcowych od pozostałej domeny 4HB dla pełnej długości domeny (WT, czerwony) bądź domeny z usuniętą helisą pierwszą (ΔH1 , zielony). Na rysunku umieszczono struktury reprezentatywne dla minimum energii swobodnej obecnych dla przypadku ΔH1 .

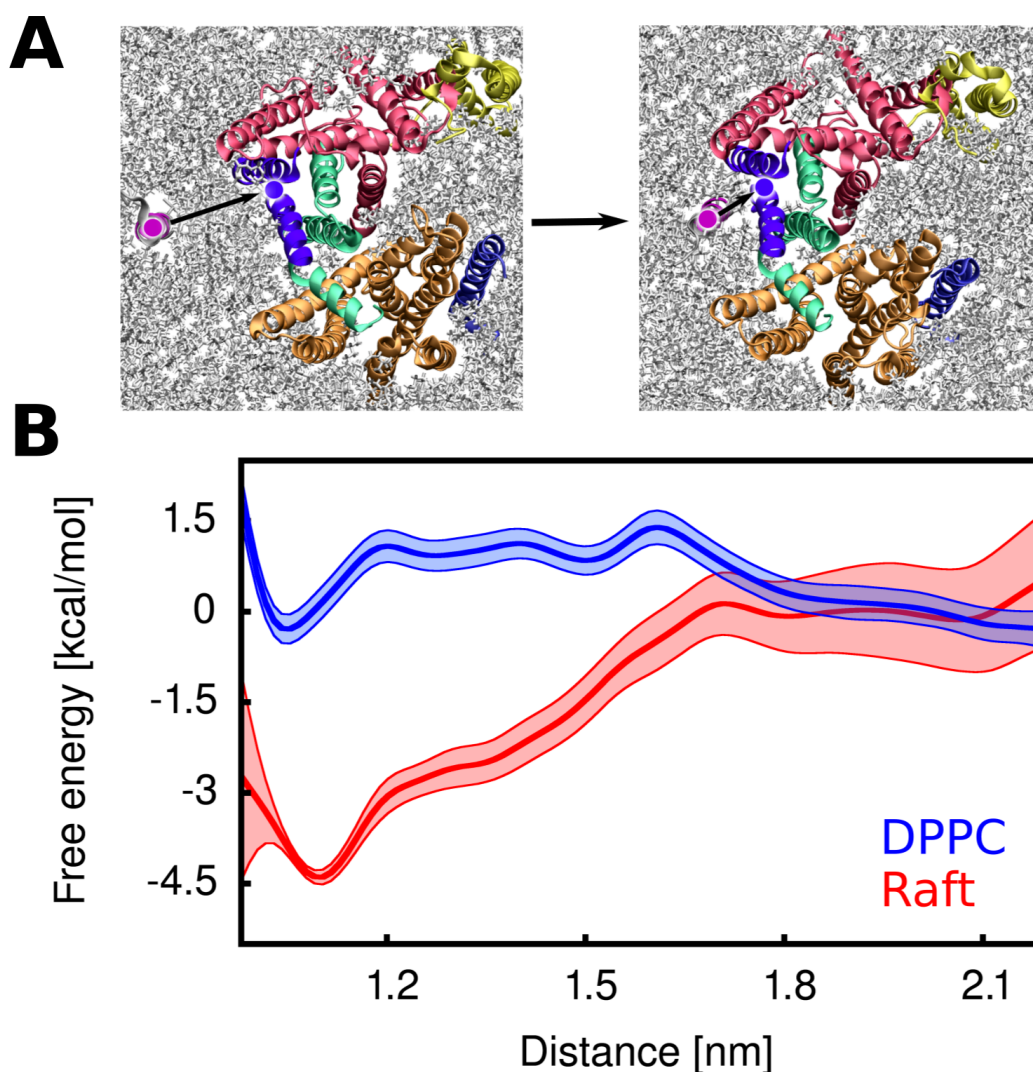
Finalnie sprawdziłem, czy zidentyfikowane zmiany w obszarze fragmentu C-końcowego domeny 4HB odpowiadają za przyjmowanie konformacji 4HB promujących aktywację czynnika transkrypcyjnego. W tym celu porównałem konformacje przyjmowane przez fragment C-końcowy dla dwóch wariantów białka – pełnego, który przyjmuje głównie konformacje zwinięte uniemożliwiające aktywację genów odpowiedzialnych za oporność wielolekowej, oraz pozbawionego helisy 1 (reszty 346–365), który przyjmuje głównie konformacje rozwinięte umożliwiające aktywację tych genów.[98] Jako współrzędną opisującą wybrałem odległość między centrum masy łańcucha głównego helisy 2 i 4 a centrum masy łańcucha głównego sześciu C-końcowych reszt domeny 4HB. Głębokie minimum energii swobodnej dla pełnego wariantu białka przy wartości współrzędnej wynoszącej około 1 nm (Rys. 4.17 świadczy o tym, że reszty C-końcowe pozostają w bliskim kontakcie z helisami 2 i 4. Usunięcie helisy 1 skutkuje pojawieniem się drugiego stanu konformacyjnego, w którym C-koniec jest całkowicie eksponowany do środowiska wodnego (Rys. 4.17). Dodatkowo, minimum odpowiadające stanowi ściśle związanemu C-końca do helisy 2 i 4 staje się płytsze oraz przesuwa się ku mniejszym wartościom (wartość współrzędnej opisującej wynoszącej około 0.8 nm), co wskazuje na zmianę powierzchni wiążącej C-koniec (tworzonej przez helisy 2 i 4), prawdopodobnie w wyniku przyjmowania luźniejszej struktury przez całą domenę. O ile więc powyższy wynik nie pozwala na jednoznaczne zakwalifikowanie zmian konformacyjnych obserwowanych dla pełnego wariantu białka w symulacjach replik temperaturowych, o tyle pokazuje, że wypętlanie się C-końca z reszty domeny 4HB pozwala na aktywację czynnika transkrypcyjnego aktywującego ekspresję genów odpowiedzialnych za oporność wielolekową przez 4HB.

4.3 Rozpoznawanie i wiązanie substratów przez γ -sekretazę

4.3.1 Cholesterol odgrywa istotną rolę w wiązaniu substratów przez γ -sekretazę

Choć γ -sekretaza przeprowadza wewnątrz błonową proteolizę przynajmniej 90 różnych substratów, to jednak nie obserwuje się między nimi analogii jeśli chodzi o sekwencje. [100, 183] Co interesujące, pokazano, że aktywność enzymu zależy w istotnym stopniu od zawartości cholesterolu w środowisku lipidowym, w którym znajduje się enzym. Wzrost zawartości cholesterolu w błonie stymuluje aktywność proteolityczną enzymu, [105] zaś jego usunięcie z błony skutkuje całkowitym zahamowaniem procesu cięcia substratów przez enzym. [104] Postanowiłem zatem sprawdzić na przykładzie domeny transbłonowej bezpośredniego prekursora polipeptydów β -amyloidowych, czy przy pomocy dynamiki molekularnej można odtworzyć obserwowany efekt obecności cholesterolu na wzrost powinowactwa γ -sekretazy do substratów. W tym celu wyznaczyłem profile energii swobodnej dla asocjacji substratu do części transbłonowej enzymu w czystej błonie DPPC oraz w błonie złożonej z DPPC i cholesterolu w stosunku 3:2, stanowiącej model domen błony wzbogaconych w cholesterol (tzw. tratw lipidowych). [184, 185] Jako współrzędną opisującą wybrałem odległość między centrami mas łańcuchów głównych substratu (reszty 706–718) oraz helisy 6 i 9 preseniliny (reszty 242–262 oraz 430–451), znajdującymi się w bezpośrednim sąsiedztwie obu potencjalnych szczelin odpowiadających za dostęp substratu do centrum aktywnego enzymu. Jak zaprezentowano

na Rys. 4.18, w przypadku błony wzbogaconej w cholesterol obserwowane jest szerokie minimum o głębokości około -3 kcal mol $^{-1}$ dla odległości około 1 nm między substratem a preseniliną, świadczące o samorzutności tworzenia kompleksu γ -sekretyazy-substrat. W przypadku czystej błony DPPC minimum energii swobodnej odpowiadające stanowi związanej jest bardzo płytkie (Rys. 4.18), a sam proces asocjacji jest niekorzystny termodynamicznie. Powyższe wyniki świadczą o tym, że cholesterol odgrywa istotną rolę w wiązaniu substratów przez γ -sekretyzę już na etapie tworzenia wstępnego kompleksu.



RYSUNEK 4.18: **A** Przedstawienie współrzędnej opisującej wstępne wiązanie substratu do γ -sekretyazy. **B** Profile energii swobodnej dla wstępnego wiązania substratu do γ -sekretyazy w błonie złożonej z DPPC (niebieska) oraz w modelu tratwy lipidowej (czerwona).

4.3.2 Błona ulega lokalnemu ściemnieniu w bezpośrednim sąsiedztwie wstępnego miejsca wiązania substratu γ -sekretyazy

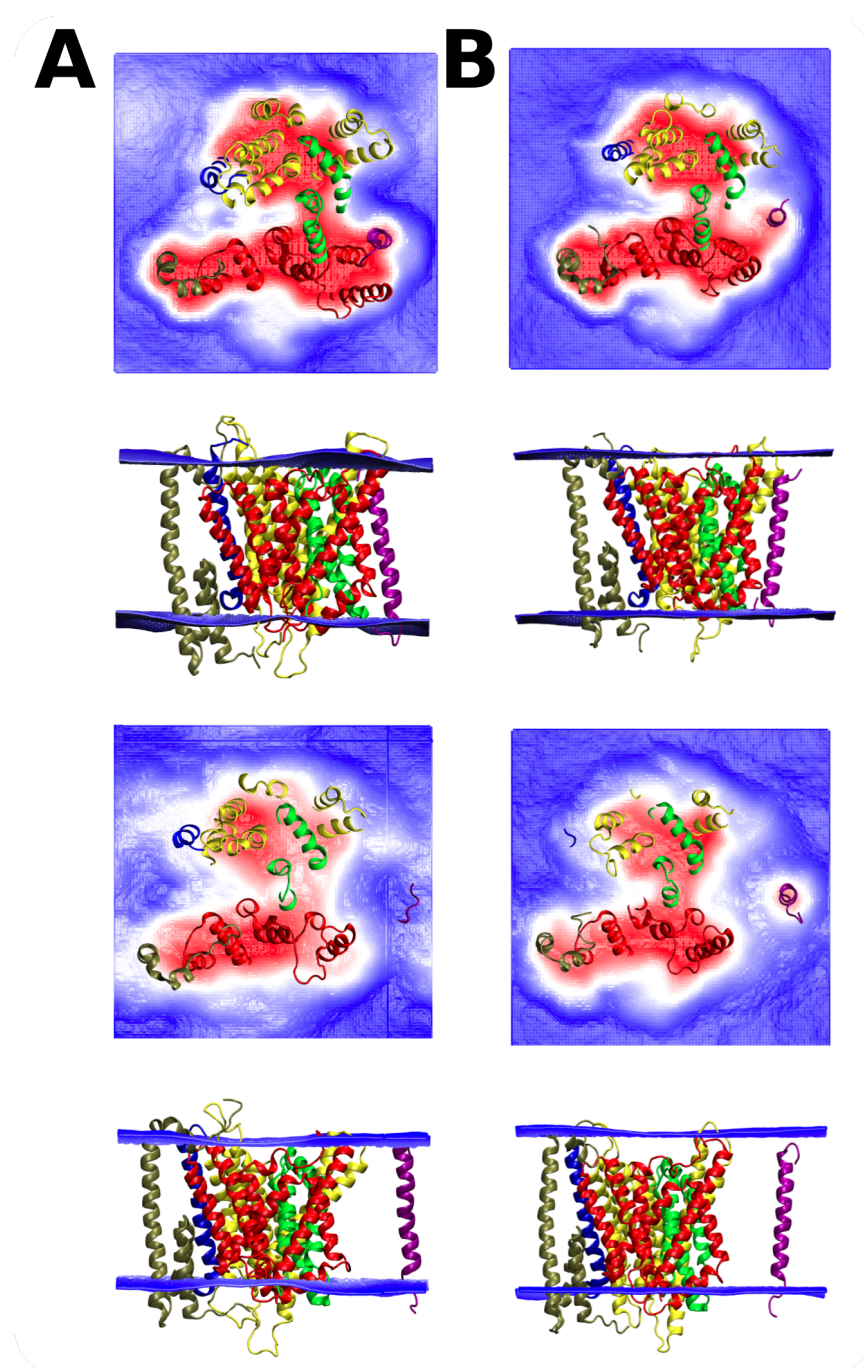
Cholesterol obecny w błonie porządkuje konformacyjnie i usztywnia sąsiadujące z nim łańcuchy lipidowe, co skutkuje m.in. zwiększeniem grubości błony. [186–188]

Występowanie różnic między długością transbłonowego fragmentu białka a grubością hydrofobową błony leży u podstaw efektu niedopasowania hydrofobowego, który może indukować agregację białek w błonie. [38] Postanowiłem zatem sprawdzić, na ile grubości błon różnią się między badanymi układami, dla których wyznaczyłem energię swobodną asocjacji substratu do γ -sekretazy. Otrzymane grubości błon, zdefiniowane jako odległość między centrami mas atomów fosforu każdego z listków względem normalnej błony, wynosiły odpowiednio (4.08 ± 0.06) nm dla błony złożonej z DPPC oraz (4.52 ± 0.04) nm dla modelu tratwy lipidowej, co daje różnicę grubości wynoszącą ponad 0.4 nm. Pozwala to podejrzewać, że niedopasowanie hydrofobowe jest jednym z efektów różnicujących oddziaływanie białek między badanymi układami.

Pragnąc dokładniej przyjrzeć się w jaki sposób substrat i γ -sekretaza wpływają na lokalne właściwości błony, wyznaczyłem lokalne grubości błony DPPC oraz modelu tratwy lipidowej w sąsiedztwie domeny transbłonowej prekursora polipeptydów amyloidowych oraz γ -sekretazy. Jak pokazano na Rys. 4.19, choć w przypadku DPPC błona jest nieznacznie cieńsza w sąsiedztwie miejsca wiązania substratu, to sam substrat znajdujący się daleko od enzymu nie wpływa na lokalne właściwości błony. Przeciwnie jest w modelu tratwy lipidowej, gdzie błona również ulega lokalnemu ścieleniu w bezpośrednim sąsiedztwie. Związanie enzymu z substratem w modelu tratwy skutkuje zatem mniejszym obszarem błony, który musi ulec zakrzywieniu w wyniku sąsiedztwa białek o krótkich fragmentach transbłonowych. Wynik ten wskazuje, że niedopasowanie hydrofobowe substratu i enzymu do błony może stanowić istotną siłę napędową dla rozpoznania i wiązania substratów przez γ -sekretazę.

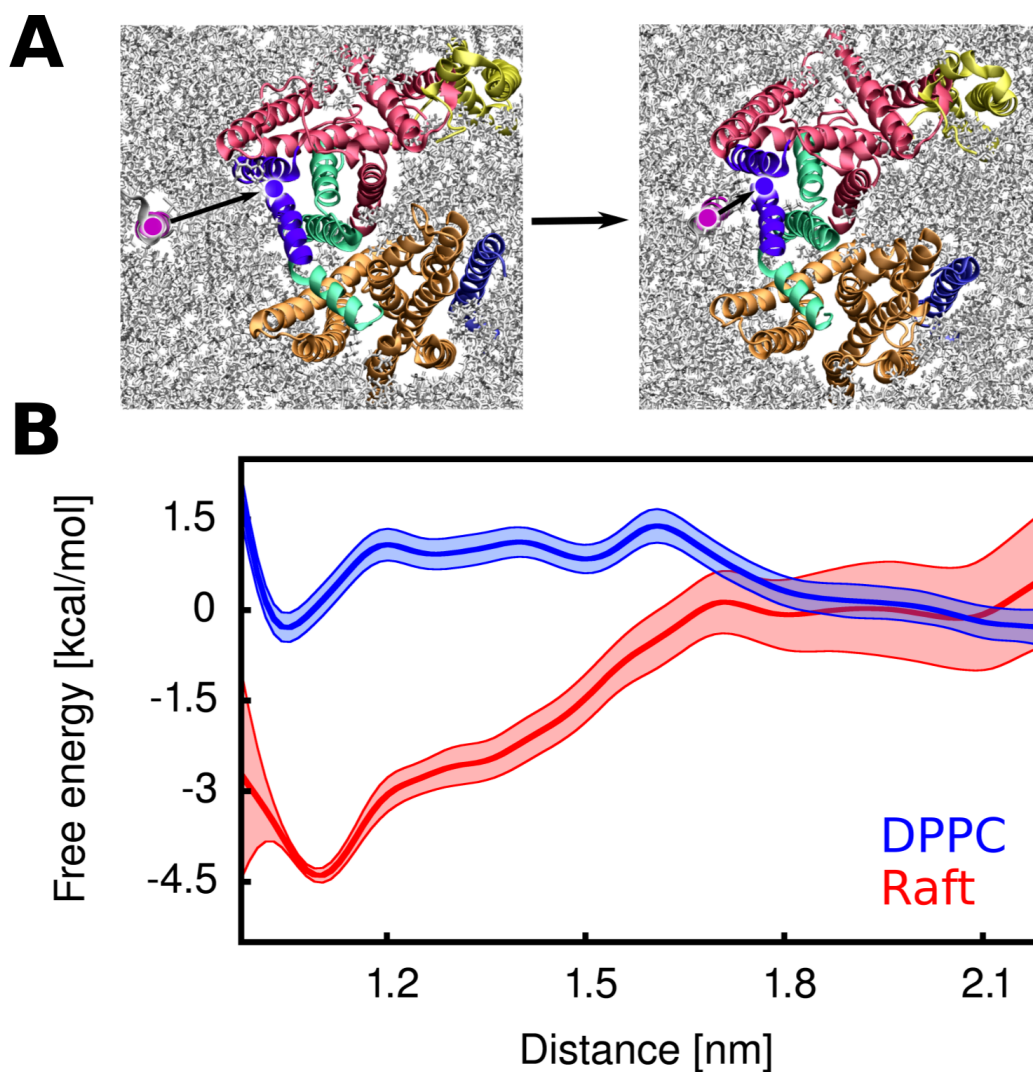
4.3.3 Efekt niedopasowania hydrofobowego stanowi siłę napędową leżącą u podstaw asocjacji substratu do wstępnego miejsca wiązania γ -sekretazy

Zachęcony uzyskanymi wcześniej wynikami, postanowiłem sprawdzić postawioną hipotezę, że efekt niedopasowania hydrofobowego odgrywa istotną rolę w wiązaniu substratów przez γ -sekretazę. W tym celu początkowo próbowałem wyznaczyć profil energii swobodnej dla tworzenia kompleksu w błonie DPPC o sztucznie zwiększonej grubości błony pomocy dodatkowego potencjału, przez co miała ona imitować grubość tratwy lipidowej. Niestety, sztuczne zwiększanie grubości błony dodatkowo ułatwiało porządkowanie lipidów, co przy wysokiej temperaturze przemiany fazowej błony DPPC (314 K) prowadziło do powstawania fazy żelowej w trakcie symulacji, uniemożliwiając uzyskanie uzbieżnionego profilu energii swobodnej. By móc oszacować na ile wzrost powinowactwa wynika z różnic grubości błony, sprawdzić ten efekt dla białek umieszczonych w błonie złożonej z 1-stearoil-2-oleilofosfatydylocholiny (SOPC). Błona SOPC ma dwie istotne zalety – z jednej strony oddaje grubość błony DPPC, [189] a z drugiej ma niską temperaturę przejścia fazowego (279.7 K), [190] co pozwala uniknąć niechcianych przejść fazowych pod wpływem sztucznego zwiększania jej grubości.



RYSUNEK 4.19: Lokalne grubości błon naokoło γ -sekreazy i substratu w kompleksie (górze) bądź wolnych (dół) dla **A** modelu tratwy lipidowej bądź **B** błony złożonej z DPPC. Kolorem czerwonym, ścienione, zaś niebieskim pogrubione obszary błony.

Jak pokazano na Rys. 4.20, minimum obecne na profilu energii swobodnej dla małych odległości substrat- γ -sekreaza wyznaczonym w błonie SOPC o grubości modelu tratwy lipidowej jest o ponad 2 kcal mol^{-1} głębsze niż w przypadku analogicznego profilu wyznaczonego w błonie SOPC o normalnej grubości. Wynik ten sugeruje, że niedopasowanie hydrofobowe rzeczywiście odgrywa istotną rolę w wiązaniu substratów przez γ -sekreazę, zaś rola cholesterolu w procesie rozpoznania substratu sprowadza się do zwiększenia grubości błony. Samorzutność tworzenia



RYSUNEK 4.20: Profile energii swobodnej dla wstępnego związania substratu do γ -sekreazy w błonie złożonej z SOPC o normalnej grubości (zielony) bądź o grubości tratwy lipidowej (fioletowy).

kompleksu w obu błonach SOPC wynika najprawdopodobniej gorszej rozpuszczalności zarówno substratu, jak i enzymu w błonach o nienasyconych łańcuchach lipidowych niż w tych zawierających nasycone lipidy, co skutkuje zwiększoną tendencją do agregacji.[191, 192]

Rozdział 5

Wnioski

W niniejszej pracy na wybranych przykładach zaprezentowałem, że metody symulacji komputerowych pozwalają na identyfikację mechanizmów rozpoznania i wiązania pomiędzy białkami. Przykład białek Hsp70 pokazał, że konwencjonalne symulacje pełnoatomowej dynamiki molekularnej, przy wykorzystaniu dodatkowych informacji pochodzących np. z analiz bioinformatycznych, mogą prowadzić do uzyskania wiarygodnej struktury kompleksów białkowych. Dodatkowe symulacje wykorzystujące techniki wzmocnionego próbkowania przestrzeni konfiguracyjnej pozwoliły mi na zasugerowanie oddziaływań elektrostatycznych jako głównych sił napędowych dla wiązania białka pomocniczego JDP Jac1 do Ssq1, co zostało następnie zweryfikowane doświadczalnie z zastosowaniem technik biochemicznych przez zespół prof. Marszałka. Dalsze badania umożliwiły mi scharakteryzowanie dynamiki pętli białka Jac1 zawierającej motyw HPD, który odgrywa istotną rolę w stymulacji cyklu katalitycznego białek Hsp70 przez białka JDP oraz na zapostulowanie mechanizmu selekcji konformacyjnej dla dokowania tej pętli do miejsca wiązania. Przy pomocy symulacji pokazałem również na przykładzie kompleksów Ssq1-Jac1 oraz DnaK-domena J białka DnaJ, jaką rolę odgrywa motyw HPD w osłabianiu oddziaływań pomiędzy domenami NBD i SBD- β białek Hsp70 w wyniku oderwania kluczowej Arg. Zaproponowałem także zmianę konformacyjną zachodzącą pod wpływem oderwania Arg od jej partnerów wiązania. Otrzymane wyniki posłużyły mi do wyjaśnienia zmian indukowanych przez mutacje zaburzające komunikację międzydomenową białek Hsp70. W końcu, przeprowadzone symulacje pozwoliły na zaproponowanie roli jaką odgrywa każdy z elementów domeny J w stymulowaniu postępu cyklu katalitycznego białek Hsp70.

Badania przeprowadzone dla domeny 4HB pozwoliły na ustalenie, w jaki sposób domena ta przystosowywała się w procesie ewolucji do pełnienia dodatkowej funkcji, jaką jest aktywowanie czynnika transkrypcyjnego odpowiedzialnego za aktywację genów oporności wielolekowej. Wykonane symulacje pokazały, że w przypadku *S. cerevisiae* stabilność domeny 4HB w przeważającej mierze wynika z efektu hydrofobowego. W kolejnym kroku pokazałem, w jaki sposób zmiany rozmieszczenia reszt tworzących hydrofobowy rdzeń wpłynęły na obniżenie stabilności domeny, co umożliwiałoby wypełnianie C-końca domeny odpowiedzialnego za indukowanie oporności.[98] Na koniec wykazałem, że zmiany w domenie 4HB zachodzące pod wpływem usunięcia helisy 1 skutkują wyeksponowaniem C-końca do wody, umożliwiając jego rozpoznanie przez czynnik transkrypcyjny.

Symulacje wykonane dla wstępnego tworzenia kompleksu z substratami przez γ -sekretazę umożliwiły mi zaproponowanie, w jaki sposób obecność cholesterolu w



blonie wpływa na wzmożoną aktywność tego enzymu. Przeprowadzone obliczenia pozwoliły nie tylko na odtworzenie efektu cholesterolu na wiązanie substratów obserwowanego eksperymentalnie, ale również na zasugerowanie efektu niedopasowania hydrofobowego jako siły napędowej leżącej u podstaw rozpoznania substratów przez ten enzym.

Całość przeprowadzonych badań pokazała, że symulacje pełnoatomowego wariantu dynamiki molekularnej stanowią narzędzie umożliwiające zarówno przewidywanie struktur kompleksów makromolekularnych, jak również identyfikację sił napędowych leżących u podstaw tworzenia kompleksów białkowych, umożliwiając im pełnienie różnorodnych funkcji biologicznych. Połączenie metod obliczeniowych omówionych w tej pracy z badaniami doświadczalnymi pozwala na rozwiązywanie problemów, które dotychczas nie mogły być skutecznie badane.

Bibliografia

1. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *eng. Science (New York, N.Y.)* **292**, 1876–1882. ISSN: 0036-8075 (czer. 2001).
2. Chen, X., Wang, Q., Ni, F. & Ma, J. Structure of the full-length Shaker potassium channel Kv1.2 by normal-mode-based X-ray crystallographic refinement. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 11352–11357. ISSN: 0027-8424 (czer. 2010).
3. Wang, W. & MacKinnon, R. Cryo-EM Structure of the Open Human Ether-à-go-go-Related K⁺ Channel hERG. *eng. Cell* **169**, 422–430.e10. ISSN: 1097-4172 (2017).
4. Mymrikov, E. V., Seit-Nebi, A. S. & Gusev, N. B. Heterooligomeric complexes of human small heat shock proteins. *Cell Stress & Chaperones* **17**, 157–169. ISSN: 1355-8145 (mar. 2012).
5. Guo, H., Suzuki, T. & Rubinstein, J. L. Structure of a bacterial ATP synthase. *eLife* **8** (red. Marletta, M. A., Berry, R. M., Mueller, D., Vik, S. & Junge, W.) e43128. ISSN: 2050-084X (lut. 2019).
6. Satagopan, S., Chan, S., Perry, L. J. & Tabita, F. R. Structure-function studies with the unique hexameric form II ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) from *Rhodospseudomonas palustris*. *eng. The Journal of Biological Chemistry* **289**, 21433–21450. ISSN: 1083-351X (sierp. 2014).
7. Nooren, I. M. A. & Thornton, J. M. Diversity of protein-protein interactions. *eng. The EMBO journal* **22**, 3486–3492. ISSN: 0261-4189 (lip. 2003).
8. Esmailbeiki, R., Krawczyk, K., Knapp, B., Nebel, J.-C. & Deane, C. M. Progress and challenges in predicting protein interfaces. *Briefings in Bioinformatics* **17**, 117–131. ISSN: 1467-5463 (sty. 2016).
9. Sikosek, T. & Chan, H. S. Biophysics of protein evolution and evolutionary protein biophysics. *eng. Journal of the Royal Society, Interface* **11**, 20140419. ISSN: 1742-5662 (list. 2014).
10. Chakrabarti, P. & Janin, J. Dissecting protein-protein recognition sites. *en. Proteins: Structure, Function, and Bioinformatics* **47**, 334–343. ISSN: 1097-0134 (2002).
11. Ma, B., Elkayam, T., Wolfson, H. & Nussinov, R. Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. *eng. Proceedings of the National Academy of Sciences of the United States of America* **100**, 5772–5777. ISSN: 0027-8424 (maj 2003).
12. Hwang, H., Petrey, D. & Honig, B. A hybrid method for protein-protein interface prediction. *Protein Science : A Publication of the Protein Society* **25**, 159–165. ISSN: 0961-8368 (sty. 2016).
13. Yan, C., Wu, F., Jernigan, R. L., Dobbs, D. & Honavar, V. Characterization of Protein-Protein Interfaces. *en. The Protein Journal* **27**, 59–70. ISSN: 1572-3887, 1573-4943 (sty. 2008).

14. Watkins, A. M., Bonneau, R. & Arora, P. S. Side-Chain Conformational Preferences Govern Protein–Protein Interactions. *Journal of the American Chemical Society* **138**, 10386–10389. ISSN: 0002-7863 (sierp. 2016).
15. Xu, D., Tsai, C. J. & Nussinov, R. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein engineering* **10**, 999–1012 (1997).
16. Fu, I., Case, D. A. & Baum, J. Dynamic water-mediated hydrogen bonding in a collagen model peptide. *Biochemistry* **54**, 6029–6037. ISSN: 0006-2960 (paź. 2015).
17. Sippel, K. H. & Quioco, F. A. Ion–dipole interactions and their functions in proteins. *Protein Science : A Publication of the Protein Society* **24**, 1040–1046. ISSN: 0961-8368 (lip. 2015).
18. Kovalev, P. V., Drozdov-Tikhomirov, L. N., Poroikov, V. V. & Alexandrov, A. A. Role of the electrostatic interactions in pre-orientation of subunits in the formation of protein-protein complexes. eng. *Journal of Biomolecular Structure & Dynamics* **22**, 111–118. ISSN: 0739-1102 (sierp. 2004).
19. Wiczór, M. & Czub, J. How proteins bind to DNA: target discrimination and dynamic sequence search by the telomeric protein TRF1. *Nucleic Acids Research* **45**, 7643–7654. ISSN: 0305-1048 (lip. 2017).
20. Skinner, J. J. *i in*. Conserved salt-bridge competition triggered by phosphorylation regulates the protein interactome. eng. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 13453–13458. ISSN: 1091-6490 (2017).
21. Tesei, G. *i in*. Self-association of a highly charged arginine-rich cell-penetrating peptide. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 11428–11433. ISSN: 0027-8424 (paź. 2017).
22. Hebert, M. J. & Russell, D. H. Hydration of Guanidinium Ions: An Experimental Search for Like-Charged Ion Pairs. eng. *The Journal of Physical Chemistry Letters* **10**, 1349–1354. ISSN: 1948-7185 (mar. 2019).
23. McGaughey, G. B., Gagné, M. & Rappé, A. K. π -Stacking Interactions ALIVE AND WELL IN PROTEINS. en. *Journal of Biological Chemistry* **273**, 15458–15463. ISSN: 0021-9258, 1083-351X (czer. 1998).
24. Lee, H., Dehez, F., Chipot, C., Lim, H.-K. & Kim, H. Enthalpy–Entropy Interplay in π -Stacking Interaction of Benzene Dimer in Water. *Journal of Chemical Theory and Computation* **15**, 1538–1545. ISSN: 1549-9618 (mar. 2019).
25. Seifert, N. A., Hazrah, A. S. & Jäger, W. The 1-Naphthol Dimer and Its Surprising Preference for π – π Stacking over Hydrogen Bonding. *The Journal of Physical Chemistry Letters* **10**, 2836–2841. ISSN: 1948-7185 (czer. 2019).
26. Grdadolnik, J., Merzel, F. & Avbelj, F. Origin of hydrophobicity and enhanced water hydrogen bond strength near purely hydrophobic solutes. en. *Proceedings of the National Academy of Sciences* **114**, 322–327. ISSN: 0027-8424, 1091-6490 (sty. 2017).
27. Kastiris, P. L. & Bonvin, A. M. J. J. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. eng. *Journal of the Royal Society, Interface* **10**, 20120835. ISSN: 1742-5662 (lut. 2013).
28. Lapenta, F., Aupič, J., Strmšek, Ž. & Jerala, R. Coiled coil protein origami: from modular design principles towards biotechnological applications. eng. *Chemical Society Reviews* **47**, 3530–3542. ISSN: 1460-4744 (maj 2018).
29. Silverman, B. D. Hydrophobicity of transmembrane proteins: Spatially profiling the distribution. *Protein Science : A Publication of the Protein Society* **12**, 586–599. ISSN: 0961-8368 (mar. 2003).



30. White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. eng. *Annual Review of Biophysics and Biomolecular Structure* **28**, 319–365. ISSN: 1056-8700 (1999).
31. Mravic, M. *i in*. Packing of apolar side chains enables accurate design of highly stable membrane proteins. en. *Science* **363**, 1418–1423. ISSN: 0036-8075, 1095-9203 (mar. 2019).
32. Arkhipov, A. *i in*. Architecture and membrane interactions of the EGF receptor. eng. *Cell* **152**, 557–569. ISSN: 1097-4172 (sty. 2013).
33. Khadria, A. S. *i in*. A Gly-zipper motif mediates homodimerization of the transmembrane domain of the mitochondrial kinase ADCK3. eng. *Journal of the American Chemical Society* **136**, 14068–14077. ISSN: 1520-5126 (paź. 2014).
34. Nadezhdin, K. D., Bocharova, O. V., Bocharov, E. V. & Arseniev, A. S. Dimeric structure of transmembrane domain of amyloid precursor protein in micellar environment. *FEBS Letters* **586**, 1687–1692. ISSN: 0014-5793 (czer. 2012).
35. Prakash, A., Janosi, L. & Doxastakis, M. GxxxG Motifs, Phenylalanine, and Cholesterol Guide the Self-Association of Transmembrane Domains of ErbB2 Receptors. *Biophysical Journal* **101**, 1949–1958. ISSN: 0006-3495 (paź. 2011).
36. Senes, A., Ubarretxena-Belandia, I. & Engelman, D. M. The C α -H \cdots O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions. en. *Proceedings of the National Academy of Sciences* **98**, 9056–9061. ISSN: 0027-8424, 1091-6490 (lip. 2001).
37. Anderson, S. M., Mueller, B. K., Lange, E. J. & Senes, A. Combination of C α -H Hydrogen Bonds and van der Waals Packing Modulates the Stability of GxxxG-Mediated Dimers in Membranes. eng. *Journal of the American Chemical Society* **139**, 15774–15783. ISSN: 1520-5126 (2017).
38. Killian, J. A. Hydrophobic mismatch between proteins and lipids in membranes. eng. *Biochimica Et Biophysica Acta* **1376**, 401–415. ISSN: 0006-3002 (list. 1998).
39. Botelho, A. V., Huber, T., Sakmar, T. P. & Brown, M. F. Curvature and hydrophobic forces drive oligomerization and modulate activity of rhodopsin in membranes. eng. *Biophysical Journal* **91**, 4464–4477. ISSN: 0006-3495 (grud. 2006).
40. Parton, D. L., Klingelhoefer, J. W. & Sansom, M. S. Aggregation of Model Membrane Proteins, Modulated by Hydrophobic Mismatch, Membrane Curvature, and Protein Class. *Biophysical Journal* **101**, 691–699. ISSN: 0006-3495 (sierp. 2011).
41. De Meyer, F. J.-M., Venturoli, M. & Smit, B. Molecular Simulations of Lipid-Mediated Protein-Protein Interactions. *Biophysical Journal* **95**, 1851–1865. ISSN: 0006-3495 (sierp. 2008).
42. Miura, K. An Overview of Current Methods to Confirm Protein-Protein Interactions. *Protein and Peptide Letters* **25**, 728–733. ISSN: 0929-8665 (sierp. 2018).
43. Brückner, A., Polge, C., Lentze, N., Auerbach, D. & Schlattner, U. Yeast Two-Hybrid, a Powerful Tool for Systems Biology. *International Journal of Molecular Sciences* **10**, 2763–2788. ISSN: 1422-0067 (czer. 2009).
44. Lin, J.-S. & Lai, E.-M. Protein-Protein Interactions: Co-Immunoprecipitation. eng. *Methods in Molecular Biology (Clifton, N.J.)* **1615**, 211–219. ISSN: 1940-6029 (2017).
45. Louche, A., Salcedo, S. P. & Bigot, S. Protein-Protein Interactions: Pull-Down Assays. eng. *Methods in Molecular Biology (Clifton, N.J.)* **1615**, 247–255. ISSN: 1940-6029 (2017).

46. Mahmood, T. & Yang, P.-C. Western Blot: Technique, Theory, and Trouble Shooting. *North American Journal of Medical Sciences* **4**, 429–434. ISSN: 2250-1541 (wrz. 2012).
47. Wu, Y., Li, Q. & Chen, X.-Z. Detecting protein-protein interactions by Far western blotting. eng. *Nature Protocols* **2**, 3278–3284. ISSN: 1750-2799 (2007).
48. Smits, A. H. & Vermeulen, M. Characterizing Protein-Protein Interactions Using Mass Spectrometry: Challenges and Opportunities. eng. *Trends in Biotechnology* **34**, 825–834. ISSN: 1879-3096 (2016).
49. Lorber, B., Fischer, F., Bailly, M., Roy, H. & Kern, D. Protein analysis by dynamic light scattering: methods and techniques for students. eng. *Biochemistry and Molecular Biology Education: A Bimonthly Publication of the International Union of Biochemistry and Molecular Biology* **40**, 372–382. ISSN: 1539-3429 (grud. 2012).
50. Korasick, D. A. & Tanner, J. J. Determination of protein oligomeric structure from small-angle X-ray scattering. eng. *Protein Science: A Publication of the Protein Society* **27**, 814–824. ISSN: 1469-896X (2018).
51. Tang, X. & Bruce, J. E. Chemical cross-linking for protein-protein interaction studies. eng. *Methods in Molecular Biology (Clifton, N.J.)* **492**, 283–293. ISSN: 1064-3745 (2009).
52. Krishnamurthy, M. *i in*. Caught in the act: covalent crosslinking captures activator-coactivator interactions in vivo. *ACS chemical biology* **6**, 1321–1326. ISSN: 1554-8929 (grud. 2011).
53. Yang, B. *i in*. Spontaneous and specific chemical cross-linking in live cells to capture and identify protein interactions. En. *Nature Communications* **8**, 2240. ISSN: 2041-1723 (grud. 2017).
54. Broom, H. R. *i in*. Combined Isothermal Titration and Differential Scanning Calorimetry Define Three-State Thermodynamics of fALS-Associated Mutant Apo SOD1 Dimers and an Increased Population of Folded Monomer. *Biochemistry* **55**, 519–533. ISSN: 0006-2960 (sty. 2016).
55. Pierce, M. M., Raman, C. S. & Nall, B. T. Isothermal titration calorimetry of protein-protein interactions. eng. *Methods (San Diego, Calif.)* **19**, 213–221. ISSN: 1046-2023 (paź. 1999).
56. Douzi, B. Protein-Protein Interactions: Surface Plasmon Resonance. eng. *Methods in Molecular Biology (Clifton, N.J.)* **1615**, 257–275. ISSN: 1940-6029 (2017).
57. Margineanu, A. *i in*. Screening for protein-protein interactions using Förster resonance energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM). en. *Scientific Reports* **6**, 28186. ISSN: 2045-2322 (czer. 2016).
58. Chodera, J. D. & Mobley, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. eng. *Annual Review of Biophysics* **42**, 121–142. ISSN: 1936-1238 (2013).
59. Yang, H., Yang, S., Kong, J., Dong, A. & Yu, S. Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. en. *Nature Protocols* **10**, 382–396. ISSN: 1750-2799 (mar. 2015).
60. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols* **1**, 2876–2890. ISSN: 1754-2189 (2006).
61. Cavalli, A., Salvatella, X., Dobson, C. M. & Vendruscolo, M. Protein structure determination from NMR chemical shifts. en. *Proceedings of the National Academy of Sciences* **104**, 9615–9620. ISSN: 0027-8424, 1091-6490 (czer. 2007).
62. Parker, M. Protein Structure from X-Ray Diffraction. *Journal of Biological Physics* **29**, 341–362. ISSN: 0092-0606 (grud. 2003).



63. Lyumkis, D. Challenges and opportunities in cryo-EM single-particle analysis. en. *Journal of Biological Chemistry* **294**, 5181–5197. ISSN: 0021-9258, 1083-351X (mar. 2019).
64. Kiraly, P. *i in*. Real-time pure shift 15N HSQC of proteins: a real improvement in resolution and sensitivity. *Journal of Biomolecular Nmr* **62**, 43–52. ISSN: 0925-2738 (2015).
65. DeLisle, C. F., Mendis, H. B. & Lorieu, J. L. Super resolution NOESY spectra of proteins. eng. *Journal of biomolecular NMR* **73**, 105–116. ISSN: 1573-5001 (kw. 2019).
66. Frueh, D. P., Goodrich, A., Mishra, S. & Nichols, S. NMR methods for structural studies of large monomeric and multimeric proteins. *Current opinion in structural biology* **23**, 734–739. ISSN: 0959-440X (paź. 2013).
67. Tenboer, J. *i in*. Time-Resolved Serial Crystallography Captures High Resolution Intermediates of Photoactive Yellow Protein. *Science (New York, N.Y.)* **346**, 1242–1246. ISSN: 0036-8075 (grud. 2014).
68. Ishigami, I. *i in*. Crystal structure of CO-bound cytochrome c oxidase determined by serial femtosecond X-ray crystallography at room temperature. eng. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 8011–8016. ISSN: 1091-6490 (2017).
69. Søndergaard, C. R., Garrett, A. E., Carstensen, T., Pollastri, G. & Nielsen, J. E. Structural Artifacts in Protein–Ligand X-ray Structures: Implications for the Development of Docking Scoring Functions. *Journal of Medicinal Chemistry* **52**, 5673–5684. ISSN: 0022-2623 (wrz. 2009).
70. Singh, D., Berntsen, K. R. M., Baakman, C., Vriend, G. & Lahiri, T. A Critical Note on Symmetry Contact Artifacts and the Evaluation of the Quality of Homology Models. en. *Symmetry* **10**, 25 (sty. 2018).
71. Gewering, T., Janulienė, D., Ries, A. B. & Moeller, A. Know your detergents: A case study on detergent background in negative stain electron microscopy. eng. *Journal of Structural Biology* **203**, 242–246. ISSN: 1095-8657 (wrz. 2018).
72. Bonomi, M. & Vendruscolo, M. Determination of protein structural ensembles using cryo-electron microscopy. *Current Opinion in Structural Biology. SI: 56 Sequences and topology (2019)* **56**, 37–45. ISSN: 0959-440X (czer. 2019).
73. Cressey, D. & Callaway, E. Cryo-electron microscopy wins chemistry Nobel. en. *Nature News* **550**, 167 (paź. 2017).
74. Denning, P. J. & Lewis, T. G. Exponential Laws of Computing Growth. *Commun. ACM* **60**, 54–65. ISSN: 0001-0782 (grud. 2016).
75. Song, Y. *i in*. High resolution comparative modeling with RosettaCM. *Structure (London, England : 1993)* **21**. ISSN: 0969-2126. doi:10.1016/j.str.2013.08.005 (paź. 2013).
76. Yang, J. & Zhang, Y. Protein Structure and Function Prediction Using I-TASSER. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.]* **52**, 5.8.1–5.8.15. ISSN: 1934-3396 (grud. 2015).
77. Czub, J., Wieczór, M., Prokopowicz, B. & Grubmüller, H. Mechanochemical Energy Transduction during the Main Rotary Step in the Synthesis Cycle of F1-ATPase. eng. *Journal of the American Chemical Society* **139**, 4025–4034. ISSN: 1520-5126 (2017).
78. Palermo, G., Miao, Y., Walker, R. C., Jinek, M. & McCammon, J. A. CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations. en. *Proceedings of the National Academy of Sciences* **114**, 7260–7265. ISSN: 0027-8424, 1091-6490 (lip. 2017).

79. Lane, T. J., Shukla, D., Beauchamp, K. A. & Pande, V. S. To Milliseconds and Beyond: Challenges in the Simulation of Protein Folding. *Current opinion in structural biology* **23**, 58–65. ISSN: 0959-440X (lut. 2013).
80. Sborgi, L. *i in*. Interaction Networks in Protein Folding via Atomic-Resolution Experiments and Long-Time-Scale Molecular Dynamics Simulations. *Journal of the American Chemical Society* **137**, 6506–6516. ISSN: 0002-7863 (maj 2015).
81. Miao, Y., Feixas, F., Eun, C. & McCammon, J. A. Accelerated molecular dynamics simulations of protein folding. *en. Journal of Computational Chemistry* **36**, 1536–1549. ISSN: 1096-987X (2015).
82. Deplazes, E., Jayatilaka, D. & Corry, B. Testing the use of molecular dynamics to simulate fluorophore motions and FRET. *eng. Physical chemistry chemical physics: PCCP* **13**, 11045–11054. ISSN: 1463-9084 (czer. 2011).
83. Banerjee, R., Jayaraj, G. G., Peter, J. J., Kumar, V. & Mapa, K. Monitoring conformational heterogeneity of the lid of DnaK substrate-binding domain during its chaperone cycle. *en. The FEBS Journal* **283**, 2853–2868. ISSN: 1742-4658 (2016).
84. Rico, F., Russek, A., González, L., Grubmüller, H. & Scheuring, S. Heterogeneous and rate-dependent streptavidin–biotin unbinding revealed by high-speed force spectroscopy and atomistic simulations. *en. Proceedings of the National Academy of Sciences* **116**, 6594–6601. ISSN: 0027-8424, 1091-6490 (kw. 2019).
85. Plattner, N., Doerr, S., De Fabritiis, G. & Noé, F. Complete protein–protein association kinetics in atomic detail revealed by molecular dynamics simulations and Markov modelling. *en. Nature Chemistry* **9**, 1005–1011. ISSN: 1755-4349 (paź. 2017).
86. Pan, A. C. *i in*. Atomic-level characterization of protein–protein association. *en. Proceedings of the National Academy of Sciences* **116**, 4244–4249. ISSN: 0027-8424, 1091-6490 (mar. 2019).
87. Han, W. & Schulten, K. Fibril Elongation by A β 17–42: Kinetic Network Analysis of Hybrid-Resolution Molecular Dynamics Simulations. *Journal of the American Chemical Society* **136**, 12450–12460. ISSN: 0002-7863 (wrz. 2014).
88. Wassenaar, T. A. *i in*. High-Throughput Simulations of Dimer and Trimer Assembly of Membrane Proteins. The DAFT Approach. *Journal of Chemical Theory and Computation* **11**, 2278–2291. ISSN: 1549-9618 (maj 2015).
89. Malinverni, D., Jost Lopez, A., De Los Rios, P., Hummer, G. & Barducci, A. Modeling Hsp70/Hsp40 interaction by multi-scale molecular simulations and coevolutionary sequence analysis. *eLife* **6** (red. Brunger, A. T.) e23471. ISSN: 2050-084X (maj 2017).
90. Fan, C.-Y., Lee, S. & Cyr, D. M. Mechanisms for regulation of Hsp70 function by Hsp40. *Cell Stress & Chaperones* **8**, 309–316. ISSN: 1355-8145 (paź. 2003).
91. Daugaard, M., Rohde, M. & Jäättelä, M. The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *FEBS Letters. Cellular Stress* **581**, 3702–3710. ISSN: 0014-5793 (lip. 2007).
92. Mayer, M. P. & Bukau, B. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences* **62**, 670–684. ISSN: 1420-682X (mar. 2005).
93. Stetz, G. & Verkhivker, G. M. Computational Analysis of Residue Interaction Networks and Coevolutionary Relationships in the Hsp70 Chaperones: A Community-Hopping Model of Allosteric Regulation and Communication. *eng. PLoS computational biology* **13**, e1005299. ISSN: 1553-7358 (2017).

94. Suh, W.-C. *i in*. Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. en. *Proceedings of the National Academy of Sciences* **95**, 15223–15228. ISSN: 0027-8424, 1091-6490 (grud. 1998).
95. Kityk, R., Kopp, J. & Mayer, M. P. Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. eng. *Molecular Cell* **69**, 227–237.e4. ISSN: 1097-4164 (2018).
96. Lee, K., Sharma, R., Shrestha, O. K., Bingman, C. A. & Craig, E. A. Dual interaction of the Hsp70 J-protein cochaperone Zuotin with the 40S and 60S ribosomal subunits. en. *Nature Structural & Molecular Biology* **23**, 1003–1010. ISSN: 1545-9985 (list. 2016).
97. Eisenman, H. C. & Craig, E. A. Activation of pleiotropic drug resistance by the J-protein and Hsp70-related proteins, Zuo1 and Ssz1. eng. *Molecular Microbiology* **53**, 335–344. ISSN: 0950-382X (lip. 2004).
98. Ducett, J. K. *i in*. Unfolding of the C-Terminal Domain of the J-Protein Zuo1 Releases Autoinhibition and Activates Pdr1-Dependent Transcription. *Journal of Molecular Biology* **425**, 19–31. ISSN: 0022-2836 (sty. 2013).
99. Brown, M. S., Ye, J., Rawson, R. B. & Goldstein, J. L. Regulated Intramembrane Proteolysis: A Control Mechanism Conserved from Bacteria to Humans. English. *Cell* **100**, 391–398. ISSN: 0092-8674, 1097-4172 (lut. 2000).
100. Haapasalo, A. & Kovacs, D. M. The Many Substrates of Presenilin/ γ -Secretase. *Journal of Alzheimer's Disease* **25**, 3–28. ISSN: 1387-2877 (2011).
101. Yuan, X. *i in*. Notch signaling: An emerging therapeutic target for cancer treatment. *Cancer Letters* **369**, 20–27. ISSN: 0304-3835 (grud. 2015).
102. Boumenir, A. *i in*. CSF level of β -amyloid peptide predicts mortality in Alzheimer's disease. *Alzheimer's Research & Therapy* **11**, 29. ISSN: 1758-9193 (mar. 2019).
103. Bai, X.-c. *i in*. An atomic structure of human γ -secretase. *Nature* **525**, 212–217. ISSN: 0028-0836 (wrz. 2015).
104. Wahrle, S. *i in*. Cholesterol-dependent gamma-secretase activity in buoyant cholesterol-rich membrane microdomains. eng. *Neurobiology of Disease* **9**, 11–23. ISSN: 0969-9961 (lut. 2002).
105. Osenkowski, P., Ye, W., Wang, R., Wolfe, M. S. & Selkoe, D. J. Direct and potent regulation of gamma-secretase by its lipid microenvironment. eng. *The Journal of Biological Chemistry* **283**, 22529–22540. ISSN: 0021-9258 (sierp. 2008).
106. Wolozin, B. Statins and therapy of Alzheimer's disease: questions of efficacy versus trial design. *Alzheimer's Research & Therapy* **3**, 3. ISSN: 1758-9193 (sty. 2012).
107. Beck, F. X., Grünbein, R., Lugmayr, K. & Neuhofer, W. Heat shock proteins and the cellular response to osmotic stress. eng. *Cellular Physiology and Biochemistry: International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology* **10**, 303–306. ISSN: 1015-8987 (2000).
108. Young, J. C. Mechanisms of the Hsp70 chaperone system. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **88**, 291–300. ISSN: 0829-8211 (kw. 2010).
109. Voos, W. Chaperone-protease networks in mitochondrial protein homeostasis. eng. *Biochimica Et Biophysica Acta* **1833**, 388–399. ISSN: 0006-3002 (lut. 2013).
110. Liu, Q., Krzewska, J., Liberek, K. & Craig, E. A. Mitochondrial Hsp70 Ssc1: role in protein folding. eng. *The Journal of Biological Chemistry* **276**, 6112–6118. ISSN: 0021-9258 (mar. 2001).

111. Zhang, P., Leu, J. I.-J., Murphy, M. E., George, D. L. & Marmorstein, R. Crystal Structure of the Stress-Inducible Human Heat Shock Protein 70 Substrate-Binding Domain in Complex with Peptide Substrate. en. *PLOS ONE* **9**, e103518. ISSN: 1932-6203 (lip. 2014).
112. Bennett, W. S. & Steitz, T. A. Structure of a complex between yeast hexokinase A and glucose: II. Detailed comparisons of conformation and active site configuration with the native hexokinase B monomer and dimer. *Journal of Molecular Biology* **140**, 211–230. ISSN: 0022-2836 (czer. 1980).
113. Kityk, R., Kopp, J., Sinning, I. & Mayer, M. P. Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. eng. *Molecular Cell* **48**, 863–874. ISSN: 1097-4164 (grud. 2012).
114. Arakawa, A., Handa, N., Shirouzu, M. & Yokoyama, S. Biochemical and structural studies on the high affinity of Hsp70 for ADP. eng. *Protein Science: A Publication of the Protein Society* **20**, 1367–1379. ISSN: 1469-896X (sierp. 2011).
115. Bertelsen, E. B., Chang, L., Gestwicki, J. E. & Zuiderweg, E. R. P. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. eng. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 8471–8476. ISSN: 1091-6490 (maj 2009).
116. Brehme, M. & Voisine, C. Model systems of protein-misfolding diseases reveal chaperone modifiers of proteotoxicity. eng. *Disease Models & Mechanisms* **9**, 823–838. ISSN: 1754-8411 (2016).
117. Vickery, L. E. & Cupp-Vickery, J. R. Molecular chaperones HscA/Ssq1 and HscB/Jac1 and their roles in iron-sulfur protein maturation. eng. *Critical Reviews in Biochemistry and Molecular Biology* **42**, 95–111. ISSN: 1040-9238 (kw. 2007).
118. Bracher, A. & Verghese, J. The nucleotide exchange factors of Hsp70 molecular chaperones. *Frontiers in Molecular Biosciences* **2**. ISSN: 2296-889X. doi:10.3389/fmolb.2015.00010 (kw. 2015).
119. Wall, D., Zylicz, M. & Georgopoulos, C. The NH₂-terminal 108 amino acids of the Escherichia coli DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. eng. *The Journal of Biological Chemistry* **269**, 5446–5451. ISSN: 0021-9258 (lut. 1994).
120. Delewski, W. *i in*. Iron-Sulfur Cluster Biogenesis Chaperones: Evidence for Emergence of Mutational Robustness of a Highly Specific Protein-Protein Interaction. eng. *Molecular Biology and Evolution* **33**, 643–656. ISSN: 1537-1719 (mar. 2016).
121. Willmund, F. *i in*. The cotranslational function of ribosome-associated Hsp70 in eukaryotic protein homeostasis. eng. *Cell* **152**, 196–209. ISSN: 1097-4172 (sty. 2013).
122. Gautschi, M. *i in*. RAC, a stable ribosome-associated complex in yeast formed by the DnaK-DnaJ homologs Ssz1p and zuotin. eng. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 3762–3767. ISSN: 0027-8424 (mar. 2001).
123. Zhang, Y. *i in*. Structural basis for interaction of a cotranslational chaperone with the eukaryotic ribosome. eng. *Nature Structural & Molecular Biology* **21**, 1042–1046. ISSN: 1545-9985 (grud. 2014).
124. Peisker, K. *i in*. Ribosome-associated Complex Binds to Ribosomes in Close Proximity of Rpl31 at the Exit of the Polypeptide Tunnel in Yeast. *Molecular Biology of the Cell* **19**, 5279–5288. ISSN: 1059-1524 (grud. 2008).



125. Leidig, C. *i in.* Structural characterization of a eukaryotic chaperone–the ribosome-associated complex. eng. *Nature Structural & Molecular Biology* **20**, 23–28. ISSN: 1545-9985 (sty. 2013).
126. Braun, E. L. & Grotewold, E. Fungal Zuotin proteins evolved from MIDA1-like factors by lineage-specific loss of MYB domains. eng. *Molecular Biology and Evolution* **18**, 1401–1412. ISSN: 0737-4038 (lip. 2001).
127. Chen, D.-H., Huang, Y., Liu, C., Ruan, Y. & Shen, W.-H. Functional conservation and divergence of J-domain-containing ZUO1/ZRF orthologs throughout evolution. en. *Planta* **239**, 1159–1173. ISSN: 1432-2048 (czer. 2014).
128. Boyer, L. A., Latek, R. R. & Peterson, C. L. The SANT domain: a unique histone-tail-binding module? eng. *Nature Reviews. Molecular Cell Biology* **5**, 158–163. ISSN: 1471-0072 (lut. 2004).
129. Hundley, H. A., Walter, W., Bairstow, S. & Craig, E. A. Human Mpp11 J protein: ribosome-tethered molecular chaperones are ubiquitous. eng. *Science (New York, N.Y.)* **308**, 1032–1034. ISSN: 1095-9203 (maj 2005).
130. Richly, H. *i in.* Transcriptional activation of polycomb-repressed genes by ZRF1. eng. *Nature* **468**, 1124–1128. ISSN: 1476-4687 (grud. 2010).
131. Shrestha, O. K. *i in.* Structure and evolution of the 4-helix bundle domain of Zuotin, a J-domain protein co-chaperone of Hsp70. eng. *PloS One* **14**, e0217098. ISSN: 1932-6203 (2019).
132. Jurisch-Yaksi, N., Sannerud, R. & Annaert, W. A fast growing spectrum of biological functions of γ -secretase in development and disease. eng. *Biochimica Et Biophysica Acta* **1828**, 2815–2827. ISSN: 0006-3002 (grud. 2013).
133. Brookmeyer, R., Gray, S. & Kawas, C. Projections of Alzheimer’s disease in the United States and the public health impact of delaying disease onset. eng. *American Journal of Public Health* **88**, 1337–1342. ISSN: 0090-0036 (wrz. 1998).
134. VanDussen, K. L. *i in.* Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells. eng. *Development (Cambridge, England)* **139**, 488–497. ISSN: 1477-9129 (lut. 2012).
135. Li, X. *i in.* Structure of a presenilin family intramembrane aspartate protease. en. *Nature* **493**, 56–61. ISSN: 1476-4687 (sty. 2013).
136. Kornilova, A. Y., Bihel, F., Das, C. & Wolfe, M. S. The initial substrate-binding site of gamma-secretase is located on presenilin near the active site. eng. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3230–3235. ISSN: 0027-8424 (mar. 2005).
137. Li, Y. *i in.* Structural interactions between inhibitor and substrate docking sites give insight into mechanisms of human PS1 complexes. eng. *Structure (London, England: 1993)* **22**, 125–135. ISSN: 1878-4186 (sty. 2014).
138. Sato, C., Takagi, S., Tomita, T. & Iwatsubo, T. The C-terminal PAL motif and transmembrane domain 9 of presenilin 1 are involved in the formation of the catalytic pore of the gamma-secretase. eng. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **28**, 6264–6271. ISSN: 1529-2401 (czer. 2008).
139. Bai, X.-c., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human γ -secretase. *eLife* **4** (red. Kühlbrandt, W.) e11182. ISSN: 2050-084X (grud. 2015).
140. Yang, G. *i in.* Structural basis of Notch recognition by human γ -secretase. En. *Nature* **565**, 192. ISSN: 1476-4687 (sty. 2019).
141. Zhou, R. *i in.* Recognition of the amyloid precursor protein by human γ -secretase. en. *Science* **363**, eaaw0930. ISSN: 0036-8075, 1095-9203 (lut. 2019).

142. Takasugi, N. *i in.* The role of presenilin cofactors in the gamma-secretase complex. eng. *Nature* **422**, 438–441. ISSN: 0028-0836 (mar. 2003).
143. Xie, T. *i in.* Crystal structure of the γ -secretase component nicastrin. en. *Proceedings of the National Academy of Sciences* **111**, 13349–13354. ISSN: 0027-8424, 1091-6490 (wrz. 2014).
144. Shah, S. *i in.* Nicastrin functions as a gamma-secretase-substrate receptor. eng. *Cell* **122**, 435–447. ISSN: 0092-8674 (sierp. 2005).
145. Bolduc, D. M., Montagna, D. R., Gu, Y., Selkoe, D. J. & Wolfe, M. S. Nicastrin functions to sterically hinder γ -secretase–substrate interactions driven by substrate transmembrane domain. en. *Proceedings of the National Academy of Sciences* **113**, E509–E518. ISSN: 0027-8424, 1091-6490 (lut. 2016).
146. Barrett, P. J. *i in.* The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. eng. *Science (New York, N.Y.)* **336**, 1168–1171. ISSN: 1095-9203 (czer. 2012).
147. Deatherage, C. L., Lu, Z., Kim, J.-H. & Sanders, C. R. Notch Transmembrane Domain: Secondary Structure and Topology. eng. *Biochemistry* **54**, 3565–3568. ISSN: 1520-4995 (czer. 2015).
148. Landry, S. J. Structure and energetics of an allele-specific genetic interaction between dnaJ and dnaK: correlation of nuclear magnetic resonance chemical shift perturbations in the J-domain of Hsp40/DnaJ with binding affinity for the ATPase domain of Hsp70/DnaK. eng. *Biochemistry* **42**, 4926–4936. ISSN: 0006-2960 (maj 2003).
149. Barnett, R. N. & Landman, U. Born-Oppenheimer molecular-dynamics simulations of finite systems: Structure and dynamics of $(\text{H})_2$. *Physical Review B* **48**, 2081–2097 (lip. 1993).
150. Zhou, Y., Wang, S., Li, Y. & Zhang, Y. Born-Oppenheimer Ab Initio QM/MM Molecular Dynamics Simulations of Enzyme Reactions. eng. *Methods in Enzymology* **577**, 105–118. ISSN: 1557-7988 (2016).
151. Noé, F. Beating the Millisecond Barrier in Molecular Dynamics Simulations. *Biophysical Journal* **108**, 228–229. ISSN: 0006-3495 (sty. 2015).
152. Perilla, J. R. & Schulten, K. Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations. en. *Nature Communications* **8**, 15959. ISSN: 2041-1723 (lip. 2017).
153. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics* **98**, 10089–10092. ISSN: 0021-9606 (czer. 1993).
154. Liu, P., Kim, B., Friesner, R. A. & Berne, B. J. Replica exchange with solute tempering: A method for sampling biological systems in explicit water. en. *Proceedings of the National Academy of Sciences* **102**, 13749–13754. ISSN: 0027-8424, 1091-6490 (wrz. 2005).
155. Den Otter, W. K. Revisiting the Exact Relation between Potential of Mean Force and Free-Energy Profile. *Journal of Chemical Theory and Computation* **9**, 3861–3865. ISSN: 1549-9618 (wrz. 2013).
156. Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. en. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 826–843. ISSN: 1759-0884 (2011).
157. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. en. *Journal of Computational Chemistry* **13**, 1011–1021. ISSN: 1096-987X (1992).

158. Van Der Spoel, D. *i in.* GROMACS: fast, flexible, and free. eng. *Journal of Computational Chemistry* **26**, 1701–1718. ISSN: 0192-8651 (grud. 2005).
159. Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. eng. *Journal of Computational Chemistry* **34**, 2135–2145. ISSN: 1096-987X (wrz. 2013).
160. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926–935. ISSN: 0021-9606 (lip. 1983).
161. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. eng. *Biophysical Journal* **97**, 50–58. ISSN: 1542-0086 (lip. 2009).
162. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **81**, 511–519. ISSN: 0021-9606 (lip. 1984).
163. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182–7190. ISSN: 0021-8979 (grud. 1981).
164. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation* **4**, 116–122. ISSN: 1549-9618 (sty. 2008).
165. Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* **159**, 98–103 (lip. 1967).
166. Torrie, G. M. & Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* **23**, 187–199. ISSN: 0021-9991 (lut. 1977).
167. Laio, A. & Parrinello, M. Escaping free-energy minima. en. *Proceedings of the National Academy of Sciences* **99**, 12562–12566. ISSN: 0027-8424, 1091-6490 (paź. 2002).
168. Bonomi, M. *i in.* PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **180**, 1961–1972. ISSN: 0010-4655 (paź. 2009).
169. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38. ISSN: 0263-7855 (lut. 1996).
170. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. en. *Proceedings of the National Academy of Sciences* **98**, 10037–10041. ISSN: 0027-8424, 1091-6490 (sierp. 2001).
171. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. eng. *Nucleic Acids Research* **32**, W665–667. ISSN: 1362-4962 (lip. 2004).
172. Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. eng. *Physical Review Letters* **100**, 020603. ISSN: 0031-9007 (sty. 2008).
173. Piana, S. & Laio, A. A bias-exchange approach to protein folding. eng. *The Journal of Physical Chemistry. B* **111**, 4553–4559. ISSN: 1520-6106 (maj 2007).
174. Harrigan, M. P. *i in.* MSMBuilder: Statistical Models for Biomolecular Dynamics. eng. *Biophysical Journal* **112**, 10–15. ISSN: 1542-0086 (sty. 2017).
175. Raiteri, P., Laio, A., Gervasio, F. L., Micheletti, C. & Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *The Journal of Physical Chemistry B* **110**, 3533–3539. ISSN: 1520-6106 (mar. 2006).



176. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* **4**, 435–447. ISSN: 1549-9618 (mar. 2008).
177. Hitzenberger, M. & Zacharias, M. γ -Secretase Studied by Atomistic Molecular Dynamics Simulations: Global Dynamics, Enzyme Activation, Water Distribution and Lipid Binding. eng. *Frontiers in Chemistry* **6**, 640. ISSN: 2296-2646 (2018).
178. Gapsys, V., de Groot, B. L. & Briones, R. Computational analysis of local membrane properties. *Journal of Computer-Aided Molecular Design* **27**, 845–858. ISSN: 0920-654X (2013).
179. Ciesielski, S. J. *i in*. Interaction of J-protein co-chaperone Jac1 with Fe-S scaffold Isu is indispensable in vivo and conserved in evolution. eng. *Journal of Molecular Biology* **417**, 1–12. ISSN: 1089-8638 (mar. 2012).
180. English, C. A., Sherman, W., Meng, W. & Gierasch, L. M. The Hsp70 interdomain linker is a dynamic switch that enables allosteric communication between two structured domains. en. *Journal of Biological Chemistry* **292**, 14765–14774. ISSN: 0021-9258, 1083-351X (wrz. 2017).
181. Vogel, M., Mayer, M. P. & Bukau, B. Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. eng. *The Journal of Biological Chemistry* **281**, 38705–38711. ISSN: 0021-9258 (grud. 2006).
182. Hennessy, F., Cheetham, M. E., Dirr, H. W. & Blatch, G. L. Analysis of the levels of conservation of the J domain among the various types of DnaJ-like proteins. *Cell Stress & Chaperones* **5**, 347–358. ISSN: 1355-8145 (paź. 2000).
183. Weihofen, A. & Martoglio, B. Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides. eng. *Trends in Cell Biology* **13**, 71–78. ISSN: 0962-8924 (lut. 2003).
184. Levental, I. & Veatch, S. The continuing mystery of lipid rafts. *Journal of molecular biology* **428**, 4749–4764. ISSN: 0022-2836 (grud. 2016).
185. Risselada, H. J. & Marrink, S. J. The molecular face of lipid rafts in model membranes. en. *Proceedings of the National Academy of Sciences* **105**, 17367–17372. ISSN: 0027-8424, 1091-6490 (list. 2008).
186. Nezil, F. A. & Bloom, M. Combined influence of cholesterol and synthetic amphiphilic peptides upon bilayer thickness in model membranes. *Biophysical Journal* **61**, 1176–1183. ISSN: 0006-3495 (maj 1992).
187. Warschawski, D. E. & Devaux, P. F. Order parameters of unsaturated phospholipids in membranes and the effect of cholesterol: a ^1H - ^{13}C solid-state NMR study at natural abundance. eng. *European biophysics journal: EBJ* **34**, 987–996. ISSN: 0175-7571 (list. 2005).
188. Hung, W.-C., Lee, M.-T., Chen, F.-Y. & Huang, H. W. The Condensing Effect of Cholesterol in Lipid Bilayers. *Biophysical Journal* **92**, 3960–3967. ISSN: 0006-3495 (czer. 2007).
189. Kučerka, N., Nieh, M.-P. & Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. eng. *Biochimica Et Biophysica Acta* **1808**, 2761–2771. ISSN: 0006-3002 (list. 2011).
190. Kaneshina, S., Ichimori, H., Hata, T. & Matsuki, H. Barotropic phase transitions of dioleoylphosphatidylcholine and stearoyl-oleoylphosphatidylcholine bilayer membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1374**, 1–8. ISSN: 0005-2736 (wrz. 1998).

191. Hur, J.-Y. *i in.* Active gamma-secretase is localized to detergent-resistant membranes in human brain. eng. *The FEBS journal* **275**, 1174–1187. ISSN: 1742-464X (mar. 2008).
192. Sun, F. *i in.* Dimerization and Structural Stability of Amyloid Precursor Proteins Affected by the Membrane Microenvironments. *Journal of Chemical Information and Modeling* **57**, 1375–1387. ISSN: 1549-9596 (czer. 2017).

Spis rysunków

1.1	Struktura polimerazy II RNA w kompleksie z RNA	2
1.2	Budowa białek rodziny Hsp70	10
1.3	Cykl katalityczny białka Hsp70	11
1.4	Struktura 4HB organizmów drożdżowych oraz <i>Homo sapiens</i>	13
1.5	Struktura kompleksu γ -sekretazy umieszczonego w błonie lipidowej	14
1.6	Budowa preseniliny	15
1.7	Struktura nikastryny	16
3.1	Schematyczne przedstawienie dwuwymiarowego pudła zawierającego 10 cząsteczek gazu doskonałego	21
4.1	Izopowierzchnie potencjału elektrostatycznego naokoło białka Jac1 i Ssq1 oraz schematyczne przedstawienie wstępnego ułożenia białek Ssq1 i Jac1	41
4.2	Profile energii swobodnej dla tworzenia kompleksu przez Ssq1 z Jac1 oraz jego mutantem	42
4.3	Izopowierzchnie potencjału elektrostatycznego naokoło Jac1 oraz jego mutantu	43
4.4	Reprezentatywne struktury kompleksu Jac1-Ssq1 dla największych klastrów	44
4.5	Nałożenie struktur kompleksów Jac1-Ssq1 oraz DnaJ-DnaK	45
4.6	Mapy energii swobodnej zrzutowane na płaszczyznę największej zmienności konformacyjnej domeny J	46
4.7	Reprezentatywna struktura białka Ssq1 oraz profile energii swobodnej dla zmiany partnerów oddziaływania kluczowej Arg	47
4.8	Współrzędna reakcji dla rozsunęcia domen NBD i SBD- β oraz profile energii swobodnej i rozkłady prawdopodobieństwa	49
4.9	Definicje odległości między subdomenami NBD a SBD- β oraz rozkłady prawdopodobieństwa odległości między subdomenami NBD a SBD- β	51
4.10	Schematyczne przedstawienie kąta torsyjnego między subdomenami NBD oraz rozkłady prawdopodobieństwa kąta torsyjnego między subdomenami NBD	52
4.11	Schematyczne przedstawienie porównywanych układów oraz profile energii swobodnej dla rozejścia się domen NBD i SBD- β Ssq1 bądź DnaK w stanie wolnym bez SBD- α , bądź z SBD- α w kompleksie z białkiem JDP	53
4.12	Profile energii swobodnej dla rozejścia domen NBD i SBD- β dla białek Ssq1 bądź DnaK oraz wybranych mutantów.	54
4.13	Profile energii swobodnej dla rozejścia domen NBD i SBD- β dla białka DnaK w stanie wolnym bądź w kompleksie z mutantem Asp35Ala białka J	54

4.14 Schematyczne przedstawienie podziału domeny J białka DnaJ oraz profile energii swobodnej dla rozejścia domen NBD i SBD β dla DnaK w stanie wolnym, w obecności H2 oraz Δ H3	55
4.15 Filogeneza domeny 4HB u <i>Saccharomycetaceae</i> , <i>Candida</i> , <i>Pezizomycotina</i> , <i>Animalia</i> oraz pierwotniaków	57
4.16 Schematyczne przedstawienie współrzędnej opisującej proces rozwijania domeny 4HB oraz profile przykładowej siły względem rozciągnięcia	59
4.17 Reprezentatywne struktury domeny 4HB <i>S. cerevisiae</i> dla dwóch głównych klastrów oraz profile energii swobodnej dla wyplecenia fragmentu 6 reszt C-końcowych od pozostałej domeny 4HB	60
4.18 Przedstawienie współrzędnej opisującej wstępne związanie substratu do γ -sekreazy oraz profile energii swobodnej dla wstępnego związania substratu do γ -sekreazy	62
4.19 Lokalne grubości błon naokoło γ -sekreazy i substratu	64
4.20 Profile energii swobodnej dla wstępnego związania substratu do γ -sekreazy	65

Spis tablic

3.1	Położenia centrów potencjału oraz odpowiadające im stałe siłowe symulacji umbrella sampling dla asocjacji białka Jac1 do Ssq1.	33
3.2	Definicje fragmentów domeny 4HB dla struktur z bazy PDB.	37
3.3	Definicje fragmentów domeny 4HB dla modeli homologicznych.	37
4.1	Wkłady do powierzchni tworzącej rdzeń hydrofobowy znanych struktur domeny 4HB pochodzące od jej wybranych fragmentów.	56
4.2	Wkłady do powierzchni tworzącej rdzeń hydrofobowy dla struktur domeny 4HB na linii filogenetycznej <i>S. cerevisiae</i> pochodzące od jej wybranych fragmentów.	58