

Warszawa, 23 maja 2016

Dr Joanna Ida Sułkowska

Interdyscyplinarne Laboratorium Modelowania Układów Biologicznych
Centrum Nowych Technologii

Zakład Chemii Teoretycznej i Krystalografii
Wydział Chemii

Uniwersytet Warszawski

AUTOREFERAT

Spis treści

1. IMIĘ I NAZWISKO	2
2. POSIADANE DYPLOMY I STOPNIE NAUKOWE.....	2
3. INFORMACJE O DOTYCHCZASOWYM ZATRUDNIENIU W JEDNOSTKACH NAUKOWYCH	3
4. WSKAZANIE OSIĄGNIĘCIA WYNIKAJĄCEGO Z ART. 16 UST. 2 USTAWY Z DNIA 14 MARCA 2003 R. O STOPNIACH NAUKOWYCH I TYTULE NAUKOWYM ORAZ O STOPNIACH I TYTULE W ZAKRESIE SZTUKI (DZ. U. NR 65, POZ. 595 ZE ZM.)	3
A) TYTUŁ OSIĄGNIĘCIA NAUKOWEGO	3
B) WYKAZ PUBLIKACJI STANOWIĄCYCH OSIĄGNIĘCIE NAUKOWE	3
C) OMÓWIENIE CELU NAUKOWEGO WW. PUBLIKACJI I OSIĄGNIĘTYCH WYNIKÓW	6
5. OMÓWIENIE POZOSTAŁYCH OSIĄGNIĘĆ NAUKOWO-BADAWCZYCH.....	26
A) GRANTY I PROJEKTY BADAWCZE	26
B) NAGRODY	29
C) DANE BIBLIOMETRYCZNE	30
D) WYKAZ I OMÓWIENIE POZOSTAŁYCH PUBLIKACJI	30

1. Imię i nazwisko

Joanna Ida Sułkowska

2. Posiadane dyplomy i stopnie naukowe

1. Stopień naukowy: doktor nauk fizycznych

Nadany przez: Instytut Fizyki, Polska Akademia Nauk

Rok uzyskania: 2007 (z wyróżnieniem)

Tytuł rozprawy doktorskiej: „Stretching and folding proteins in coarse grained models”

Promotor: Prof. Marek Cieplak (Instytut Fizyki, Polska Akademia Nauk)

2. Dyplom magisterski

Nadany przez: Wydział Fizyki, Uniwersytet Warszawski

Rok uzyskania: 2003

Promotorzy:

- Prof. Christoph F. Schmidt (Vrije Universiteit Amsterdam)
- Prof. Marek Cieplak (Instytut Fizyki, Polska Akademia Nauk)

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych

1. **Uniwersytet Warszawski**, Centrum Nowych Technologii
Kierownik grupy badawczej „Interdisciplinary Laboratory of Biological Systems Modelling”,
od 2014 r.
2. **Uniwersytet Warszawski**, Wydział Chemii
Adiunkt, od 2013 r.
Pracownik naukowo-techniczny, w latach 2012-2013
3. **Rice University, Houston**, Wydział Chemii
Visiting faculty, w latach 2012-2013
4. **Uniwersytet Kalifornijski San Diego**, Center for Theoretical Biological Physics
Staż doktorski, w latach 2008-2012
5. **Uniwersytet Kalifornijski San Diego**, Wydział Biochemii
Visiting faculty, w latach 2012-2013
6. **Polska Akademia Nauk**, Instytut Fizyki
Adiunkt, w latach 2008-2010

4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. nr 65, poz. 595 ze zm.)

A) Tytuł osiągnięcia naukowego

**Klasyfikacja i krajobraz energetyczny zapętlonych białek:
węzły, slipknoty oraz lassa**

B) Wykaz publikacji stanowiących osiągnięcie naukowe

H1. P. Dąbrowski-Tumanski, W. Niemyska, P. Pasznik, J.I. Sulikowska,
“*LassoProt: server to analyze biopolymers with lassos*”
Nucleic Acids Res. (2016), doi: 10.1093/nar/gkw308.
Internetowa baza danych i serwer: <http://LassoProt.cent.uw.edu.pl>

- H2. P. Dąbrowski-Tumanski, A.I. Jarmolińska, J.I. Sulkowska,
“Prediction of the optimal set of contacts to fold the smallest knotted protein”,
 J. Phys. Cond. Mat. (2015) 27(35):354109.
- H3. M. Jamroz, W. Niemyska, E.J. Rawdon, A. Stasiak, K.C. Millett, P. Sulkowski, J.I. Sulkowska,
“KnotProt: a database of proteins with knots and slipknots”,
 Nucleic Acids Res. (2014), 43: D306-D314.
 Internetowa baza danych i serwer: <http://KnotProt.cent.uw.edu.pl>
- H4. E. Haglund, J.I. Sulkowska, J.K. Noel, H. Lammert, J.N. Onuchic, P.A. Jennings,
“Pierced Lasso Bundles are a New Class of Knot Motifs”,
 PLoS Comput. Biology (2014) 19,10(6):e1003613.
- H5. P. Dąbrowski-Tumanski, S. Niewieczczal, J.I. Sulkowska,
“Determining Critical Amino Acid contacts for knotted protein”,
 TASK Quarterly (2014) 18 No 3, 323–337.
- H6. E.J. Rawdon, K.C. Millett, J.I. Sulkowska, A. Stasiak,
“Knot localization in proteins”,
 Biochemical Society Transactions (2013) 41(2):538-41.
- H7. K.C. Millett, E.J. Rawdon, A. Stasiak, J.I. Sulkowska,
“Identifying knots in proteins”,
 Biochemical Society Transactions (2013) 41(2):533-7.
- H8. T. Andrews, D.T. Capraro, J.I. Sulkowska, J.N. Onuchic, P.A. Jennings,
“Hysteresis as a Marker for Complex, Overlapping Landscapes in GFP and Other Proteins”,
 J. Phys. Chemistry Letters (2012) Jan 3;4(1):180-188.
- H9. J.K. Noel, J.N. Onuchic, J.I. Sulkowska,
“Knotting in explicit solvent”,
 J. Phys. Chem. Letters (2013) 4(21), 3570-3573.
- H10. J.I. Sulkowska, J.K. Noel, C. A. Ramírez-Sarmiento, E.J. Rawdon, K.C. Millett, J. N. Onuchic,
“Knotting pathways in proteins”,
 Biochemical Society Transactions (2013) 41(2):523-7.
- H11. J.I. Sulkowska*, J. K Noel*, J. N. Onuchic,
“Energy landscape of knotted protein folding”,
 Proc. Natl. Acad. Sci. USA 109 (2012) 17783-17788.
- H12. J.I. Sulkowska*, E. J. Rawdon*, K. C. Millett, J. N. Onuchic, A. Stasiak,
“Conservation of complex knotting and slipknotting patterns in proteins”,
 Proc. Natl. Acad. Sci. USA 109(26): E1715-23.
- H13. E. Haglund, J.I. Sulkowska, Zhao He, Gen-Sheng Feng, P. Jennings, J. N. Onuchic,
“The unique cysteine knot regulates the pleiotropic hormone leptin”,
 PLoS One (2012) 7(9) e45654.
- H14. J.K. Noel, J.I. Sulkowska, J.N. Onuchic,
“Slipknotting upon native-like loop formation in a trefoil knot protein”,
 Proc. Natl. Acad. Sci. USA 107 (2010) 15403.

H15. J.I. Sulkowska, P Sulkowski, P Szymczak, M Cieplak,
"Untying knots in proteins",
J. Am. Chem. Soc. 132 (40) (2010) 13954-13956.

H16. D. Bölinger, J.I. Sulkowska, H.P. Hsu, L.A. Mirny, M. Kardar, J.N. Onuchic, P. Virnau,
"A Stevedore's protein knot",
PLoS Comput Biol. (2010) 6, e1000731.

H17. J.I. Sulkowska, P. Sulkowski, J.N. Onuchic,
"Dodging the crisis in protein folding with knots",
Proc. Natl. Acad. Sci. USA 106 (2009) 3119-3124.

H18. J.I. Sulkowska, P. Sulkowski, J.N. Onuchic,
"Jamming proteins with slipknots and their free energy landscape",
Phys. Rev. Lett. (2009) 103 268103.

H19. J I. Sulkowska, P Sulkowski, P Szymczak, M Cieplak,
"Stabilizing effect of knots on proteins - How knots influence properties of proteins",
Proc. Natl. Acad. Sci. USA 105 (2008) 19714-19719.

C) Omówienie celu naukowego ww. publikacji i osiągniętych wyników

Spis treści

I Wstęp

II Identyfikacja złożonych topologii w białkach: węzły, spliknoty oraz lassa; klasyfikacja białek

- A Model macierzowy do opisu węzłów oraz spliknotów [H6,H7,H12]
- B Klasyfikacja węzłów i spliknotów [H3,H12]
- C Model minimalnej powierzchni do opisu lassa [H1]
- D Klasyfikacja lass [H1,H4,H13]

III Ewolucja motywów topologicznych, konfiguracja geometryczna oraz rola zapętleń w białkach

- A Zaskakujące zachowanie ewolucyjne motywów topologicznych [H12]
- B Wpływ konfiguracji geometrycznej na proces powstawania węzłów [H9-H11,H16,H17]
- C Rola węzłów [H11,H12,H16,H19]
- D Rola lass [H1,H13]

IV Krajobraz energetyczny – mechanizmy zwijania węzłów, spliknotów oraz lass

- A Mechanizmy zwijania białka z węzłem i spliknotem [H2,H4,H11,H16-17]
- B Szczegółowa analiza zwijania najmniejszego białka z węzłem [H2,H5,H9-10,H14]
- C Krajobraz energetyczny białek lassowych [H4,H13]

V Krajobraz energetyczny – rozwijanie i rozplątywanie węzłów i spliknotów

- A Zawężone białka w stanie zdenaturowanym oraz model ich rozwiązania [H8,H15,H19]
- B Unikalne metastabilne własności białek spliknotowych [H18]

I Wstęp

Białka są fundamentalnymi składnikami żywych organizmów. Zrozumienie molekularnych podstaw funkcjonowania białek wymaga poznania ich struktury i dynamiki, a także związków struktury z dynamiką. Bardzo ważną metodą badania takich zagadnień jest analiza krajobrazu energii swobodnej. Struktura około 91% poznanych białek jest jednoznacznie określona przez ich pierwszorzędową, drugorzędową i trzeciorzędową formę, a krajobraz energii swobodnej z dużym powodzeniem może być wyznaczony doświadczalnie oraz teoretycznie (np. w symulacjach komputerowych, przy użyciu modeli gruboziarnistych, a dla mniejszych biomolekuł także modeli pełnoatomowych). Jednakże badania ostatnich lat, w znacznej mierze także moje, pokazują, iż pozostałe 9% białek posiada nietrywialną topologię (tzn. są zapętlone): 1.5% takich białek posiada węzły oraz spliknoty [H3], natomiast 7.5% lassa [H1] (przykładowe węzły, spliknoty oraz lassa przedstawione są na rysunku 1). Standardowy opis konfiguracji przestrzennej (uwzględniający pierwszo-, drugo- i trzeciorzędową strukturę) takich zapętlonych białek okazuje się niewystarczający, a różne dotychczas wypracowane narzędzia i teorie wymagają istotnych modyfikacji i uwzględnienia istnienia nietrywialnej topologii.

Niniejsza praca jest poświęcona opisowi nowych, odkrytych przeze mnie zapętionych białek zawierających węzły, slipknoty oraz lassa, a także charakterystyce ich geometrii i krajobrazu energetycznego przy pomocy stworzonych przeze mnie narzędzi. Dzieło składa się z cyklu prac dotyczących przede wszystkim: wypracowania różnych metod i zastosowania ich do klasyfikacji zapętionych białek; opracowania modeli i narzędzi, zarówno teoretycznych jak i doświadczalnych, do wyznaczania własności fizyko-chemicznych kształtujących krajobraz energetyczny takich białek; określenia wpływu zapętlenia na ewolucję struktury takich białek oraz ich funkcję biologiczną. Zawartość niniejszej pracy stanowi podstawę nowej dziedziny badań, poświęconej badaniu białek o nietrywialnej topologii, która jest obecnie szeroko rozwijana na świecie.



Rys. 1 Zapętlenia, którym poświęcona jest niniejsza praca (od lewej): schemat węzła, slipknota oraz lassa.

Zanim przedstawię wyniki prowadzonych przeze mnie badań, w pierwszej kolejności zwięźle scharakteryzuję zidentyfikowane w tej pracy możliwe zapętlenia w białkach, czyli węzły, slipknoty i lassa, przedstawione schematycznie na rysunku 1.

Według definicji matematycznej za **węzeł** uważa się zamkniętą, nie przecinającą się krzywą zanurzoną w trzech wymiarach. Klasyfikacją i opisem własności węzłów zajmuje się dział matematyki zwany teorią węzłów. Jednym z narzędzi wypracowanych w ramach tej teorii są tzw. niezmienniki węzłów, czyli pewne stosunkowo proste obiekty matematyczne (np. liczby, wielomiany, itp.), które dla danego węzła można w algorytmiczny sposób wyznaczyć, i które m.in. pozwalają rozróżniać i klasyfikować węzły. Jednym z ważniejszych takich niezmienników jest wielomian Jonesa, za którego odkrycie Vaughan Jones został uhonorowany medalem Fieldsa.

Węzły odgrywają istotną rolę nie tylko w matematyce, ale także w innych dziedzinach (np. fizyce, biologii), a ich występowanie może mieć różnorakie konsekwencje (np. efektywne upakowanie DNA w kapsydzie wymaga zawężenia). Natomiast aż do ostatniej dekady XX wieku uważano, iż węzły nie występują w białkach, tzn. łańcuchy białkowe nie mogą być zawężone. Po pierwsze wynikało to z przekonania, iż powstanie węzłów w białkach byłoby energetycznie niekorzystne. Taki wniosek został sformułowany m.in. w ważnej pracy Mansfielda opublikowanej w *Nat. Struct. Biol.* (1) w 1994 r., poświęconej pierwszemu przeglądowi wyznaczonych doświadczalnie przestrzennych struktur białek, w którym zidentyfikowano tylko jedno białko z węzłem. Jednocześnie zidentyfikowany w tej pracy węzeł był płytki, tzn. posiadał tylko kilka aminokwasów (reprezentowanych przez czarne krzywe na rysunku 1), które wystają poza rdzeń węzła (niebieska krzywa). Takie płytke węzły mogą być spontanicznie rozwiązane przez fluktuacje termiczne, toteż ich istnienie uznano za mało znaczące. Ponadto analiza zawężeń w białkach wiąże się z istotnymi problemami natury matematycznej: węzły mogą być ściśle zdefiniowane tylko na krzywych zamkniętych (bez swobodnych końców), natomiast białka są krzywymi otwartymi (mają swobodne końce). W związku z tym węzły w białkach należy w odpowiedni sposób zdefiniować oraz wypracować narzędzia teoretyczne do ich analizy i klasyfikacji. Takie narzędzia, wcześniej nieznanne, wypracowane zostały przeze mnie i moich współpracowników, i przedstawione są w

dalszej części niniejszej pracy. Jak pokażę, węzły na krzywych otwartych są nie mniej intrygujące niż na krzywych zamkniętych, a ich rola w biologii może mieć fundamentalne znaczenie dla wielu procesów życiowych.

Pierwsze białko z głębokim węzłem, czyli takim, w którym co najmniej 15 aminokwasów (reprezentowanych przez czarne krzywe na rysunku 1) wystaje poza rdzeń węzła, zostało zidentyfikowane przez Taylora dopiero w 2000 r., poprzez implementację bardzo prostego algorytmu teoretycznego KMT (2). W latach 2006-2007 kolejno Grosberg i Virnau (3,4) pokazali, że gdy końce białka w konformacji natywnej są wyraźnie oddalone od rdzenia węzła, to można je jednoznacznie domknąć i wyznaczyć typ węzła (tak jak dla węzła na krzywej zamkniętej). Prace te doprowadziły do identyfikacji kilku kolejnych zapętłonych struktur. Ponadto prace te stanowiły znaczący przełom w sposobie myślenia o złożoności struktury białek, ponieważ jednoznacznie wykazały, iż istnienie węzłów w białkach jest złożonym i intrygującym zagadnieniem z pogranicza chemii, biologii, fizyki oraz matematyki, a liczba białek zawężonych – choć okazywała się coraz bardziej znacząca – jest zdecydowanie mniejsza niż należałoby się tego spodziewać na podstawie obfitości ich występowania w innych (bio)polimerach. W niniejszej pracy wyjaśnię, jaki wpływ na takie wnioski mają procesy ewolucyjnej selekcji, metody wyznaczania węzłów na krzywych otwartych, a także problemy doświadczalne z wyznaczeniem struktur białek: zawężonych, nieglobularnych, oraz nieustrukturyzowanych.

Drugą klasą zapętleń analizowanych w tej pracy są **slipknoty**, czyli obiekty złożone z dwóch pętli, z których jedna jest przewleczona przez drugą [H17-H18], tak jak pokazano na rysunku 1 (struktura środkowa). Po domknięciu końców (bez wprowadzania dodatkowych przecięć) slipknot reprezentuje konfigurację trywialną (węzeł trywialny) z punktu widzenia teorii węzłów. Natomiast jeśli z odpowiedniego końca odetnie się kilka aminokwasów, slipknot zamieni się w (otwarty) węzeł. Slipknoty okazują się posiadać bardzo interesujące własności mechaniczne [H18]. Mimo że powstanie slipknota wydaje się bardziej prawdopodobne i mniej energetycznie wymagające niż węzłów, pierwszych sześć białek o topologii slipknota zostało odkrytych dopiero 2006 roku (5). Intrygującymi pytaniami dotyczącymi slipknotów są: czy takich struktur jest więcej, jak je wyznaczyć i sklasyfikować, jak bardzo są skomplikowane, czy mają znaczenie dla mechanicznych i termodynamicznych własności białek oraz ich funkcji biologicznych. M.in. na te pytania udzielię odpowiedzi w niniejszej pracy.

Trzecią klasą zapętleń odkrytych i analizowanych w ramach niniejszej pracy są **lassa**. Zgodnie z definicją sformułowaną w pracy [H1] składają się one z kowalencyjnie zamkniętej (np. mostkiem cysteinowym, amidowym, etc.) pętli, przez którą jest przewleczony choć jeden z końców ich struktury. Najprostszy przykład lassa pokazany jest na rysunku 1 (po prawej). Inspiracją do identyfikacji i klasyfikacji tych struktur były:

- 1) własności białek z mostkami cysteinowymi, które mogą tworzyć tzw. węzły cysteinowe poprzez trzy mostki cysteinowe; analiza tych białek pokazała, iż skomplikowana geometria wprowadza szeroko rozumianą stabilność;
- 2) istnienie złożonych cząsteczek chemicznych, których lokalne połączenia, np. poprzez jony metali, w rezultacie tworzą nietrywialne topologie;
- 3) mini-białka, zbudowane z małej zamkniętej pętli (złożonej z około 12 aminokwasów), przez którą jest przewleczony jeden z końców takiego białka; konformacja ta tworzy chemiczno-topologiczny „korek” mogący blokować aktywności biologiczne białek, np. bakteryjnej polimerazy RNA;
- 4) analiza literatury i badania pokazujące, że zamknięte łańcuchy polimerów mogą tworzyć różnego rodzaju zapętlenia, np. takie jak super-skręcone (supercoiled) koliste struktury DNA;
- 5) rola teorii węzłów i splotów w zrozumieniu funkcji biologicznej topoizomeraz oraz rekombinaz, czyli dwóch klas enzymów niezbędnych dla życia, których funkcją jest kontrola topologii DNA i RNA; pionierskie w tej dziedzinie były prace dwóch matematyków Ernst'a i Summersa (1980) o tzw. „tangle calculus”, a obecnie także matematyczki M. Vazquez, która używa teorii splotów do opisu rozplątywania nowo replikowanego bakteryjnego chromosomu.

Istnienie wyżej opisanych nietrywialnych zapętleń – węzłów, slipknotów i lass – pokazuje, iż przestrzenne ułożenie łańcucha głównego białka może być dużo bardziej złożone niż dotychczas sądzono, i w szczególności może mieć ono istotny wpływ na funkcję biologiczną. Analiza struktury i roli zapętleń w białkach wymaga nowego, interdyscyplinarnego podejścia, którego przedstawieniu poświęcona jest niniejsza praca.

Istnienie zapętlnych struktur ma też istotny wpływ na metody przewidywania struktury białek oraz analizy przebiegu procesów ewolucyjnych. Zgodnie z podstawowym założeniem biologii strukturalnej sekwencja aminokwasów determinuje strukturę przestrzenną białka (7). Analiza sekwencji aminokwasów, wsparta różnymi narzędziami bioinformatycznymi, fizycznymi, statystycznymi, i chemicznymi, pozwala z większą lub mniejszą dokładnością przewidzieć strukturę białek, czego potwierdzeniem są między innymi struktury otrzymane w zawodach CASP (The Critical Assessment of Protein Structure Prediction). Odkrycie białek zapętlnych stawia jednak pod znakiem zapytania wiele metod wypracowanych do przewidywania struktury białek, jako że przy określaniu ich skuteczności nie jest brana pod uwagę poprawna topologia przewidywanej struktury. W związku z istnieniem zapętlnych struktur istotnym staje się także pytanie, jak mocno topologia (węzły, slipknoty, lassa) jest zachowana ewolucyjnie. Pierwsze prace Bujnickiego oraz Micheletti (8,9) sugerują, iż białko z topologią węzła wyewoluowało z odwęzłonego białka. Niemniej jednak prace te nie wyjaśniają czy znane do tej pory węzły przetrwały ewolucyjną selekcję, gdyż np. są kluczowe dla funkcji biologicznej białek. Wpływ topologii na ewolucję białek – czy też zależność odwrotna – jest kolejnym z zagadnień rozważanych w niniejszej pracy.

W badaniu zapętlnych białek niezwykle ważna jest analiza ich krajobrazu energetycznego, który zawiera informację m.in. o związkach pomiędzy ich strukturą i dynamiką. Krajobraz energetyczny opisuje zarówno lokalne zmiany w strukturze związane z energią aktywacji potrzebną enzymom do przeprowadzenia danej reakcji chemicznej, jak też globalne przemiany konformacyjne – zwijanie, degradację, mechaniczne rozwijanie (10) – których przebieg jest determinowany przez własności fizyko-chemiczne aminokwasów oraz geometrię układu.

W ogólności proces zwijania białek różni się od wieloetapowej chemicznej transformacji w syntezie organicznej lub pośrednich procesach metabolicznych – nie jest on determinowany specyficzną sekwencją kroków, ani nie charakteryzuje się dużą ilością stabilnych stanów przejściowych [11]. Proces zwijania białek można opisać za pomocą kilku głównych ścieżek, które tylko w przypadku bardziej złożonych białek zawierają stany przejściowe. Obecnie procesy zwijania średniej wielkości białek są opisywane najczęściej modelami uproszczonymi, natomiast modele pełnoatomowe wraz z jawnym rozpuszczalnikiem są stosowane głównie do małych białek. W konstrukcji modeli uproszczonych istnieje kilka nurtów badań, opartych albo o stan natywny białka (12,13,14), albo potencjały statystyczne (14,15,16). W niniejszej pracy przytoczę i będę weryfikowała podejście oparte na strukturze natywnej białka, którego poprawność poparta jest wszechstronnymi badaniami teoretycznymi i eksperymentalnymi. Zgodnie z tym podejściem krajobraz energetyczny białek w większości przypadków ma kształt lejka zwijania, a białka są minimalnie sfrustrowanymi heteropolimerami, mogącymi osiągnąć zwiniętą konformację podążając jedną z możliwych dróg korzystając tylko z przyciągających natywnych kontaktów (11). Zwijanie postępuje jako proces dyfuzyjny na relatywnie gładkim krajobrazie energetycznym, podczas gdy zbiór możliwych konfiguracji zwęża się wraz ze zbliżaniem się do natywnej struktury. Proces zwijania takich heteropolimerów jest z dużym sukcesem opisywany gruboziarnistymi modelami typu Go, określanymi też jako Structure Based Models (SBM), które wykorzystują informację o strukturze natywnej. Warto zaznaczyć, że opracowanie wielkoskalowych modeli teoretycznych do opisu białek i innych struktur chemicznych zostało wyróżnione nagrodą Nobla z chemii w roku 2013. Pewne wątpliwości co do słuszności tej tezy, czyli zaskakującej prostoty mechanizmów zwijania – w przeciwieństwie do postulatu, iż problem ten jest wręcz nierozwiązywalny (paradoks

Levinthal'a (17)) – pojawiają się w środowisku naukowym od co najmniej dwóch dekad. Stworzone ostatnio superkomputery, np. Anton (skonstruowany przez D. Shaw), dostarczają jednak kolejnych przekonujących argumentów za słusnością powyższej tezy. Symulacje w jawnym rozpuszczalniku pokazały, iż te same kontakty służą do zwijania białek w modelu typu SBM oraz z jawnym rozpuszczalnikiem (18). Ponadto dla większości białek liczba natywnych kontaktów jest dobrym układem współrzędnych do opisu procesu zwijania i stanu przejściowego (19). Hipoteza, iż białka są minimalnie sfrustrowanymi heteropolimerami powstałymi w procesie ewolucji, została ostatnio mocno wsparta poprzez badanie koewolucji par aminokwasów metodą Direct Coupling Analysis (DCA) (11). Koewolucja par aminokwasów pokazuje, iż mimo że minimalnie sfrustrowane sekwencje są niezmiernie rzadkie, to są one wystarczająco gęsto rozłożone w przestrzeni sekwencji, tak że połączone ścieżki mutacji aminokwasów mogą znaleźć minimalnie sfrustrowaną sekwencję o ile istnieje wystarczająca selekcyjna presja ewolucyjna (11). Ponadto założenie minimalnej frustracji jest obecnie powszechnie wykorzystywane do symulowania procesu zwijania, projektowania i przewidywania struktury białek (11). Należy jednak podkreślić, iż istnieją białka, których krajobraz energetyczny nie jest tak idealny (19).

Podsumowując, podejście oparte na strukturze natywnej białka i teorii lejka zwijania jest skuteczne i poparte wszechstronnymi badaniami. Niemniej jednak istotne staje się pytanie czy teoria lejka zwijania oraz hipoteza minimalnie sfrustrowanej sekwencji może być zastosowana do białek o nietrywialnej topologii, których przestrzeń konfiguracyjna jest ograniczona przez więzy topologiczne; ponadto ewolucyjna selekcja wymaga, by powolne zapętlanie takich białek było kompensowane poprzez nadrzędną rolę topologii dla funkcji biologicznej lub warunków w jakich białko przebywa. W niniejszej pracy sformułuję teorię zwijania i rozwijania białek z nietrywialną topologią i wykażę jej skuteczność.

Niezależnie od metod teoretycznych, oczywiście bardzo ważne jest badanie białek metodami doświadczalnymi. Obserwowany w ostatniej dekadzie niesamowity postęp w dziedzinie technik eksperymentalnych sprawił, że manipulacje przy pomocy szczypczków optycznych, czy też mikroskop sił atomowych, umożliwiają badanie pojedynczych molekuł (20,21). Ponadto zastosowanie metod komputerowych dynamiki molekularnej w modelach pełnoatomowych (22), a przede wszystkim gruboziarnistych (23), wraz z teorią Kramersa, Bell'a, Shabo, i Dudko, umożliwia obecnie interpretację danych doświadczalnych oraz ich modelowanie w szerokim zakresie krajobrazu energetycznego wraz z jego szczegółami, które często nie są dostępne w podejściu termodynamicznym. Mechaniczne rozwijanie pojedynczych molekuł jest obecnie bardzo cennym źródłem wiedzy na temat natury oraz mechanizmów działania DNA, RNA oraz białek.

Zastosowanie – także w moich pracach – metody mechanicznego rozciągania do białek pokazuje, iż geometria układu jest głównym czynnikiem determinującym proces mechanicznej denaturacji (23). Jednakże w tym przypadku również nie jest oczywiste, czy wypracowane metody i teoria będą poprawnie opisywać mechaniczne rozwijanie białek o nietrywialnej topologii, np. proces ich degradacji. Ponadto stwierdzenie, iż „pociągnięcie białka zawężonego za końce spowoduje jego zapętlenie”, generuje wiele ważnych pytań: w jaki sposób takie struktury mogą się rozwęzić, jaką rolę pełnią węzły, czy wprowadzają one dodatkową stabilność, czy obecność slipknota albo lassa także zwiększa stabilność. W niniejszej pracy przedstawię odpowiedzi na te pytania.

Powyższa dyskusja pokazuje jednoznacznie, iż 9% poznanych dotychczas struktur białek wymaga opracowania nowych metod oraz technik ich analizy. Ten cel może być zrealizowany jedynie stosując interdyscyplinarne podejście. Należy także zauważyć, iż wraz z rozwojem metod analizy białek może się okazać, iż zapętlonych struktur jest znacznie więcej, i w konsekwencji ich rola może okazać się jeszcze bardziej istotna.

W niniejszej pracy przedstawione jest interdyscyplinarne podejście łączące dziedziny fizyki, biofizyki, biochemii, oraz matematycznej teorii węzłów, do wyjaśnienia roli nietrywialnej topologii

dla funkcji białek. Opracowane pionierskie metody, mechanizmy i teorie stanowią podstawy nowo rodzącej się dziedziny badań, poświęconej badaniu białek zawężlonych, zawierających węzły, slipknoty oraz lassa. Głównymi celami i wynikami tej pracy są:

- stworzenie narzędzi do teoretycznego opisu zapętleń w białkach, odkrycie nowych struktur zawierających węzły i lassa, oraz opracowanie pierwszych na świecie klasyfikacji białek z węzłami [H6,H7,H12], slipknotami [H17,H18] oraz lassami [H1];
- wykazanie znaczenia ewolucyjnego i biologicznego zapętlonych białek [H1, H4,H11,H12,H16,H19];
- zrozumienie procesu samozawężlenia białek z głębokimi węzłami oraz slipknotami [H11,H16,H17];
- zrekonstruowanie krajobrazu energetycznego najmniejszego zawężlonego białka na podstawie symulacji pełnoatomowych i uproszczonych modeli [H2,H5,H9,H10,H14];
- zrozumienie procesu samozawężlenia białek z lassami [H4, H13] oraz scharakteryzowanie różnic w rozwijaniu białek o różnej topologii, a tym samym foldzie [H19];
- wykazanie istnienia węzła w stanie zdenaturowanym białka [H8], sformułowanie wzoru określającego prawdopodobieństwo rozwiązania węzłów [H15], oraz odkrycie i wytłumaczenie unikalnych metastabilnych stabilnych własności białek slipknotowych [H18].

Plan niniejszej pracy jest następujący. W rozdziale II przedstawione są metody analizy nietrywialnych topologii w białkach oraz odkryte przeze mnie takie struktury. W rozdziale III wykażę, że topologia węzłów i slipknotów jest zachowana ewolucyjnie oraz przedstawię znaczenie i rolę zapętlenia dla funkcji biologicznej oraz własności mechanicznych białek. W rozdziale IV scharakteryzuję mechanizm powstawania węzłów w białkach. W rozdziale V wykażę istnienie węzła w stanie zdenaturowanym, pokażę jak go rozwiązać oraz scharakteryzuję unikalne zachowanie slipknotowych białek podczas mechanicznego rozsupływania.

II Identyfikacja zapętleń w białkach: węzły, slipknoty, oraz lassa; klasyfikacja białek

W tej części opiszę odkryte przeze mnie złożone topologie w białkach oraz modele do ich identyfikacji [H1,H6,H7,H12,H18]. Ponadto przedstawię krótkie wprowadzenie do teorii węzłów oraz scharakteryzuję wcześniej odkryte zawężlone białka. Białka z węzłami z slipknotami opiszę w punktach **A** i **B**. Należy podkreślić, że o ile pewne białka z węzłami znane były już wcześniej, to zagadnienie topologii lassa, opisane w punktach **C** i **D**, jest zupełnie nowym, nie badanym wcześniej tematem.

II.A Model macierzowy do opisu węzłów oraz slipknotów

Przypomnijmy, że zgodnie z teorią węzłów za węzeł uważa się zamkniętą, nie przecinającą się krzywą zanurzoną w trzech wymiarach. Klasyfikacji węzłów dokonuje się na podstawie tzw. niezmienników, czyli obiektów które zależą tylko od topologii węzła, a nie szczegółów jego kształtu. Jeśli dla danych dwóch węzłów odpowiednie niezmienniki są różne, oznacza to, iż nie można tych węzłów nałożyć na siebie bez dokonania przecięcia. Najprostszym niezmiennikiem jest minimalna liczba skrzyżowań otrzymana po zrzutowaniu węzła na dwuwymiarową powierzchnię. Niezmiennik ten jest też podstawą powszechnie stosowanego nazewnictwa węzłów, przy czym kolejne węzły o tej samej liczbie skrzyżowań rozróżnia się dodatkowym numerem, który umieszczany jest w dolnym indeksie liczby skrzyżowań. Najprostszym węzłem jest tzw. węzeł trywialny, oznaczany 0_1 , dla którego liczba skrzyżowań wynosi zero; kolejny to trójlistnik o trzech skrzyżowaniach, oznaczany

3_1 ; kolejne to węzeł ósemkowy o czterech skrzyżowaniach, oznaczany 4_1 , dwa węzły o pięciu skrzyżowaniach 5_1 i 5_2 , węzły 6_1 , 6_2 i 6_3 o sześciu skrzyżowaniach; ilość węzłów z większą liczbą skrzyżowań zaczyna gwałtownie wzrastać. Bardziej skomplikowane niezmienniki, pozwalające na odróżnienie większej ilości węzłów, przyjmują postać np. wielomianów jednej lub kilku zmiennych, takich jak wielomian Alexandera, Jonesa lub HOMFLY. Tego typu wielomianów także używamy do identyfikacji węzłów w białkach.

Jako, że teoria węzłów opisuje konfiguracje topologiczne krzywych zamkniętych, to identyfikacja i klasyfikacja węzłów w białkach, czyli na krzywych otwartych, wymaga pewnego jej uogólnienia. Określenie rodzaju węzła w białku wymaga jego domknięcia (połączenia w jakiś sposób jego końców, tak by łańcuch białkowy utworzył krzywą zamkniętą), a sposób takiego domknięcia ma fundamentalne znaczenie dla identyfikacji powstałego typu węzła, a więc zrozumienia roli jaką on pełni z punktu widzenia globalnych oraz lokalnych (np. miejsce aktywne) własności biologicznych, termicznych i mechanicznych biomolekuł. Zastosowanie i uogólnienie metod teorii węzłów dla przypadku krzywych otwartych wymaga zatem uwzględnienia niejednoznaczności identyfikacji węzłów na otwartych łańcuchach. Ponadto szczegółowe opisanie struktury białka wymaga rozważenia zapętleń wszystkich podłańcuchów danego łańcucha peptydowego (analiza jedynie pełnej długości białka pomija możliwe wewnętrzne zapętlenia). Przytoczone tu argumenty są inspirowane wcześniejszymi pracami, zwłaszcza Grosberga, Taylora, Yeatesa, oraz własnymi rozważaniami.

W pracy [H12] opracowałam metodę określającą statystyczne zapętlenie (tzn. najbardziej prawdopodobny typ węzła) dla każdego podłańcucha danego łańcucha peptydowego. Przez podłańcuch rozumiem tu łańcuch peptydowy z pewną liczbą obciętych aminokwasów od strony grupy amidowej (tzw. N-końiec) oraz grupy karboksylowej (C-końiec). Metoda ta nazywana jest modelem macierzowym, jako że najbardziej prawdopodobne zapętlenie każdego podłańcucha można zwięźle przedstawić w postaci macierzy, zwanej macierzą zapętleń („knot fingerprint”), czy też odpowiadającego jej diagramu. Każdy element macierzy zapętleń, odpowiadający danemu podłańcuchowi, wyznaczany jest niezależnie. Model ten po pierwsze rozwiązuje problem niejednoznaczności w sposobie wyznaczenia węzłów na otwartych krzywych, a po drugie – dzięki analizie wszystkich podłańcuchów – umożliwia wykrycie i scharakteryzowanie geometrii slipknota. Macierz zapętleń pozwala też na wizualną identyfikację regionów tworzących różne typy węzłów oraz ich wzajemne geometryczne ułożenie. Szczegóły konstrukcji macierzy zapętleń, dobór metody domykania końców, oraz zalety zaproponowanego modelu są przeanalizowane w dwóch pracach przeglądowych [H6,H7].

Model macierzy ponadto pozwala na wyznaczenie krajobrazu zmian topologicznych w białku, np. podczas zawężania w procesie zwijania białka. Zoptymalizowana wersja modelu macierzowego pozwalająca na wyznaczenie krajobrazu topologicznego jest obecnie powszechnie dostępna poprzez skonstruowany przeze mnie internetowy serwer do analizy topologicznej białek *KnotProt* (<http://KnotProt.cent.uw.edu.pl>) [H3].

II.B Klasyfikacja węzłów oraz slipknotów

Jednym z istotnych rezultatów niniejszej pracy jest sporządzona przeze mnie klasyfikacja wszystkich białek z węzłami i slipknotami. Analizując wszystkie zdeponowane w bazie PDB białka przy pomocy opisanego wyżej modelu macierzowego scharakteryzowałam występujące w nich typy i aranżacje zapętleń [H9]. Na przykład wykazałam, iż podłańcuchy odkrytego przez mnie w pracy [H13] białka z węzłem 6_1 zawierają także drugi węzeł tego samego typu 6_1 , węzeł 4_1 oraz węzeł 3_1 . Węzły te ujawniają się wzdłuż łańcucha peptydowego wraz z usuwaniem kolejnych aminokwasów odpowiednio z N-końca lub C-końca lub obu końców naraz (węzeł 3_1). Innym przykładem jest nowy typ slipknota zawierający (w różnych podłańcuchach) węzły $3_1,4_1,3_1$, które także powstają po usunięciu kolejno aminokwasów z końca C (powstaje węzeł 3_1), z końca N (węzeł 4_1), lub obu

końców równocześnie (węzeł 3_1). W pracy [H9] także wykazałam (potwierdzając tym samym wcześniejsze przypuszczenia (10,31)), iż zmiana topologii białka jest możliwa wyłącznie poprzez topologię slipknota, który przeprowadza węzły typu skręconego (tzw. „twist knots”) w formę trywialną redukując jedno skrzyżowanie.

Charakterystyka zapętleń w białkach poprzez macierz zapętleń ma następujące kluczowe znaczenia: po pierwsze pokazuje, iż każde białko posiada unikalny charakter zapętlenia (ang. „fingerprint”), po drugie pozwala w przejrzysty sposób odróżnić węzły od slipknotów (przykłady będą podane poniżej), po trzecie pozwala sklasyfikować zapętleń w białkach – białka z tą samą macierzą zapętleń należą do tej samej klasy topologicznej. Po czwarte pozwala zidentyfikować powtarzalność złożonej charakterystyki zapętleń, a w konsekwencji sugeruje ewolucyjne znaczenie węzłów, które zostanie omówione w rozdziale III. Szczegółowy opis innych motywów, jak i nowe odkryte przeze mnie zawężone białka są opisane w pracy [H9].

Kolejnym ważnym wynikiem wchodzącym w cykl niniejszej pracy jest stworzenie klasyfikacji topologicznej białek [H9] oraz jej upublicznienie w postaci bazy danych *KnotProt* – pierwszej na świecie bazy internetowej podającej pełną informację o sposobie zwięźlenia białka, dostępnej pod adresem <http://KnotProt.cent.uw.edu.pl> [H2]. Szczegóły tej klasyfikacji oraz charakterystyki białek są przedstawione w pracy [H9] i pracach przeglądowych [H5,H6]. W tych pracach wyróżnionych zostało 16 klas motywów zapętleń w białkach (Tabela 1, [H9]), oznaczanych symbolem K lub S (oznaczającym odpowiednio węzeł lub slipknot, jako rodzaj zapętlenia całego łańcucha) wraz z listą wszystkich węzłów wykrytych w podłańcuchach. Dwie najliczniejsze grupy białek zapętlnych zawierają pojedynczy węzeł lub slipknot typu 3_1 . Trzecia najliczniejsza klasa zawiera motyw $S_{3_1,4,3_1}$ zidentyfikowany w białkach membranowych. Zapętleń w białkach membranowych zidentyfikowane przez *KnotProt* zasługują na szczególną uwagę – w czterech takich białkach zidentyfikowany został węzeł 3_1 (węzły w białkach membranowych nie były wcześniej znane!), natomiast w aż około 20% białkach membranowych wykryty został slipknot.

Baza *KnotProt* stanowi kompendium wiedzy na temat nietrywialnych motywów węzłowych oraz slipknotowych w białkach, oraz zawiera ich wszechstronną klasyfikację strukturalną i biologiczną. Baza *KnotProt* automatycznie uaktualnia się co tydzień (na podstawie nowo deponowanych struktur w PDB), a zatem stanowi najbardziej aktualne źródło wiedzy o węzłach i slipknotach w białkach.

Warto podkreślić, że *KnotProt* pełni także funkcję serwera do samodzielnej analizy otwartych polimerów (otrzymanych np. jako wynik dynamiki molekularnej). Serwer ten może mieć wiele różnych zastosowań, np. może służyć do sprawdzenia poprawności topologii struktur wyznaczonych eksperymentalnie (zwłaszcza techniką Cryo-electron microscopy, cryo-EM, przy której najczęściej prowadzi do niewłaściwych zapętleń), czy też wyników uzyskiwanych w ramach zawodów CASP (serwer przelicza duże zbiory danych z żądaną szybkością). *KnotProt* jest bardzo często używany do analizy zapętlenia w symulacjach komputerowych białek i polimerów (zgodnie z moją wiedzą jest jedynym powszechnie dostępnym narzędziem do badania dynamiki zapętlenia biomolekuł).

II.C Model minimalnej powierzchni do opisu topologii lasa

Kolejnym rezultatem tej pracy jest opracowanie metod identyfikacji struktur typu laso. W pracy [H1] zdefiniowałam geometrię lasa oraz opracowałam *model minimalnej powierzchni* do jej identyfikacji. Topologię lasa mają białka zawierające zamkniętą, np. mostkiem cysteinowym lub amidowym, pętlę, którą przecina choć jeden ze swobodnych końców białka. Ścisłej rzecz ujmując, taki swobodny koniec przecina wyznaczoną teoretycznie powierzchnię o minimalnym polu, rozpiętą na kowalencyjnie zamkniętej pętli. Powierzchnia o takim minimalnym polu zachowuje się podobnie jak bańka mydlana rozpięta na zamkniętej pętli, a jej kształt może być wyznaczony przy użyciu metod matematycznych opartych o techniki triangulacji. Tak skonstruowaną powierzchnię nazwałam powierzchnią minimalną. Konstrukcja powierzchni minimalnej wraz z narzędziami do określenia

kierunku przecinania, położenia przecięcia względem zamknięcia, oraz głębokości (ilości węgli $C\alpha$ nad lub pod powierzchnią) pozwala w niemal jednoznaczny sposób określić czy białko tworzy topologię lasso, oraz jego złożoność. Taka geometryczna charakterystyka przecięcia pozwala także na określenie i identyfikację płytkich lass. Płytkie lasso może się rozwinąć pod wpływem termicznych fluktuacji (analogicznie jak w przypadku płytkich węzłów). Skomplikowane konfiguracje łańcucha głównego, np. w przypadku istnienia samoprzecięć powierzchni minimalnej, są charakteryzowane poprzez analizę dodatkowej powierzchni barymetrycznej. Należy także podkreślić, iż identyfikacja typu lasso „gołym okiem” jest dużo trudniejsza niż węzłów, i niemal niemożliwa dla dużych pętli. Skonstruowany przeze mnie model jest pierwszym narzędziem na świecie (zgodnie z moją wiedzą) do identyfikacji zapętlenia typu lasso.

Należy też zauważyć, iż sformułowanie definicji zapętlenia lasso i poszukiwanie białek z topologią lasso było umotywowane pracą nad białkiem leptyną, w której po raz pierwszy zidentyfikowałam najprostszą topologię lasso [H11], a następnie w pracy [H4] skonstruowałam prototyp minimalnej powierzchni do analizy czterech kolejnych białek.

II.D Klasyfikacja białek z lassami

Jednym z głównych wyników tej pracy jest odkrycie i scharakteryzowanie topologii lasso w białkach. Przy użyciu opisanego wyżej modelu minimalnej powierzchni, analizując reprezentatywny zbiór białek z mostkami cysteinowymi (o podobieństwie sekwencyjnym nie większym niż 35%) wykazałam, iż 18% z nich – czyli zaskakująco dużo – posiada topologię lasso [H1]. Odkryte przeze mnie konfiguracje, schematycznie przedstawione na rysunku 2, sklasyfikowałam jako następujące motywy:

- pojedyncze lasso L_1 – powierzchnia minimalna rozpięta na pętli białka jest przecinana raz;
- podwójne lasso L_2 – po przecięciu powierzchni minimalnej ten sam koniec białka przecina ją ponownie w przeciwnym kierunku (w sumie powierzchnia jest przecinana dwukrotnie);
- potrójne lasso L_3 – topologia podobna do L_2 jednakże ten sam koniec białka przecina pętlę jeszcze raz w przeciwnym kierunku (w sumie powierzchnia jest przecinana trzy razy);
- superskręcone lasso LS_i – powierzchnia pętli jest przecinana co najmniej dwa razy w tym samym kierunku (konformacja ta przypomina superskręcone koliste DNA), a indeks i oznacza ilość tych przecięć;
- dwustronne lasso LL_{ij} – powierzchnia minimalna jest przecinana co najmniej raz przez każdy z końców białka, indeksy i, j oznaczają ilość przecięć odpowiednio przez amidowy (N-koniec) oraz karboksylowy koniec (C-koniec) białka.



Rysunek 2: Odkryte przeze mnie konfiguracje typu lasso w białkach (od lewej): L_1 (pojedyncze lasso), L_2 (podwójne lasso), L_3 (potrójne lasso), LS_i (superskręcone lasso, indeks i oznacza ilość przecięć pętli, w tym przykładzie $i=2$), oraz LL_{ij} (dwustronne lasso, w tym przypadku z jednym przewleczeniem dla każdego z końców, $i=1, j=1$).

Niezależnie od tych wyników, na podstawie obszernej analizy częstości występowania danego typu lasso w zależności od rodzaju struktury drugorzędowej, typu przypisanego przez klasyfikację CATH, biologicznej funkcji, oraz organizmu z którego struktura pochodzi, wykazałam kilka

istotnych korelacji [H1]. W szczególności pokazałam, iż motywy typu lasso są najczęściej tworzone przez β spinki (motyw L_2 pojawia się nawet w 95%).

Charakterystyka motywów lasso w białkach została przeze mnie udostępniona w postaci pierwszego na świecie internetowego serwera i bazy danych *LassoProt*, pod adresem <http://LassoProt.cent.uw.edu.pl> [H1]. Część bazowa *LassoProt* zawiera zapętlenia wszystkich zdeponowanych struktur białek w PDB (224282 łańcuchy) z pętlą zamkniętą wiązaniem cysteinowym, amidowym, estrowym, lub tioestrowym. Część serwerowa służy do samodzielnej analizy typu zapętlenia pojedynczego łańcucha, lub też całych zbiorów danych (np. ścieżek zwijania) otrzymanych m.in. metodami dynamiki molekularnej, wraz z graficzną reprezentacją zmian topologicznych w czasie.

Baza i serwer *LassoProt* mogą być zastosowane w różnorodny sposób. W biofizyce może być użyty np. do badania mechanizmów powstawania lasso, lub do wyznaczenia nowego układu współrzędnych do analizy mechanizmów zwężenia białek. W biologii *LassoProt* może być wykorzystany do badania ewolucji struktury białek (baza *LassoProt* zawiera wszystkie homologiczne sekwencje, wraz z przyporządkowanym im zapętleniem). W biochemii może być wykorzystany do projektowania struktur (np. w przypadku mutacji punktowych, wprowadzanych w białkach w celu zwiększenia ich stabilności, serwer służyłby do weryfikacji otrzymanych motywów). Serwer może być także przydatny do poszukiwania bardziej złożonych struktur, np. splotów w (bio)polimerach. Zastosowania *LassoProt* w innych dziedzinach są przedstawione na jego stronie internetowej. Warto zauważyć, iż *LassoProt* został odwiedzony ponad 7000 razy przez użytkowników z 24 krajów w przeciągu pierwszych 3 miesięcy istnienia.

III Ewolucja motywów topologicznych, konfiguracja geometryczna oraz rola zapętleń w białkach

W tym rozdziale, w punkcie **A**, omówię zaskakujące zachowanie ewolucyjnie zawężonych motywów w białkach, natomiast w punkcie **B** wskażę ich znaczenie dla mechanizmów zwijania białek zawężonych. Następnie przedstawię i przedyskutuję znaczenie nietrywialnej topologii dla funkcji biologicznej białek zawężonych oraz lasso, odpowiednio w punktach **C** i **D**.

III.A Zaskakujące zachowanie ewolucyjne motywów topologicznych

Macierz zapętleń (omówiona w części II.A powyżej) okazuje się być charakterystyką białka, która umożliwia znalezienie korelacji pomiędzy jego sekwencją, strukturą, funkcją, oraz ewolucją. Z przeglądu struktur białek (zdeponowanych w PDB) przy użyciu macierzy zapętleń wynika, iż białka o bardzo niskim podobieństwie sekwencyjnym posiadają ten sam motyw zapętlenia. W pracy [H12] wykazałam, iż białka rozdzielone o ponad miliard lat ewolucji posiadają ten sam topologiczny motyw. Np. struktury z rodziny UCH reprezentowane przez białka ludzkie, jednokomórkowe pierwotniaki pasożytnicze, oraz drożdżowe, których podobieństwo sekwencyjne jest niższe niż 25%, posiadają ten sam topologiczny motyw $K5_2,3_1,3_1$ (ich podłańcuchy zawierają węzeł 5_2 oraz dwa różne węzły typu 3_1). Kolejnym przykładem są białka membranowe, w ogólności odpowiedzialne za transport jonów, których sekwencyjne podobieństwo jest na poziomie 9%, a wszystkie posiadają ten sam motyw zawężenia $S3_1,4_1,3_1$. Ponieważ mechanizm zwijania białka do struktury zawężonej jest wolniejszy i mniej efektywny niż ten dla białek odwężlonych o tej samej długości oraz podobnej sekwencji, zachowanie kształtu macierzy zapętleń implikuje pozytywną rolę nietrywialnej topologii dla organizmów w których te białka występują. Zależność ta sugeruje, iż motywy węzłowe

przetrwaly ewolucyjną selekcję, gdyż ich topologia przynosi korzyści dla funkcji białka lub warunków w jakich ono przebywa. Odkrycie to jest kluczowe dla rozwoju tej dziedziny badań.

Zanim przejdę do omawiania znaczenia biologicznego zawężenia warto zwrócić uwagę na fakt, iż przedstawione w pracach [H3, H12] dane pokazują, iż nawet najprostszy rodzaj węzła jest ściśle zachowany w całym klanie (tzn. grupie białek spokrewnionych poprzez podobieństwo sekwencyjne, profil HMM („hidden Markov model”) dostępny w bazie Pfam, lub strukturalnie). Moje badania pokazują, iż białka z P-wartością (ang. P-value) około 0.04 posiadają nadal tę samą topologię, natomiast przyjmuje się, iż dopiero $P < 0.001$ oznacza znaczące podobieństwo struktur. Jedynym znanym dziś wyjątkiem są białka z rodziny N-acetylorbithine transcarbamylyse (ATCases) i ornithine transcarbamylyse (OTCases), niemniej jednak uważam, iż ten wyjątek potwierdza przedstawioną powyżej regułę. Zaskakujące zachowanie motywów węzłowych sugeruje, iż topologia jest piątym, nadrzędnym elementem upakowania przestrzennego białek. Uwzględnienie motywu zapętlenia powinno w znacznym stopniu usprawnić przewidywanie przestrzennej struktury białek (np. podczas zawodów CASP), zwłaszcza gdy podobieństwo sekwencyjne różnych struktur jest bardzo niskie, ale istnieją przesłanki, iż pochodzą one z tej samej rodziny.

III.B Wpływ konfiguracji geometrycznej na proces powstawania węzłów

Model macierzowy (omówiony w części II.A) charakteryzujący zapętlenia w białkach [H12] pozwala także zidentyfikować zależność pomiędzy konfiguracją geometryczną, funkcją biologiczną, a mechanizmem powstawania węzłów. Mimo że do tej pory nie wykryto zależności pomiędzy występowaniem zapętlenia a sekwencją grupy aminokwasów, w pracy [H9] pokazałam, iż granice węzła wyznaczone modelem macierzowym są skorelowane z bardzo silnym zachowaniem sekwencyjnym glicyny (jednego z najslabiej zachowanych ewolucyjnie aminokwasów), alaniny oraz proliny. Regiony te odpowiadają miejscom zawiasowym, które zidentyfikowałam jako kluczowe dla zawiązania białka na podstawie teoretycznych rozważań, a także na podstawie symulacji numerycznych modelem gruboziarnistym [H16,H17,H18].

Model macierzowy dostarcza także istotnych informacji potrzebnych do dalszych prac nad ewolucją białek zawężonych. Jednym z możliwych podejść jest rozwijana także przeze mnie metoda na styku bioinformatyki i fizyki statystycznej, służąca do wyznaczenia koewolucji par aminokwasów. W tym podejściu model macierzowy wskazuje przestrzenie w strukturze białka w których następuje zmiana topologii, i których istnienie powinno wynikać z koewolucji par aminokwasów (24). Analiza tych elementów powinna doprowadzić do wykrycia par aminokwasów (nie tylko pojedynczych aminokwasów [H12]) odpowiedzialnych za zachowanie nietrywialnej topologii, tym samym identyfikując fizyczne oddziaływania odpowiedzialne za proces zawężania. Mechanizmy związania białek z węzłami omówię w rozdziale IV w punktach A i B.

III.C Rola węzłów w białkach

Jednym z kluczowych zagadnień dotyczących badanych w tej pracy zapętleń jest ich wpływ na funkcje białek. Korelacje pomiędzy nietrywialną topologią a funkcją biologiczną lub warunkami (chemicznymi, termicznymi) w jakich funkcjonuje białko widać obecnie w kilku przykładach, które omówiłam w pracach [H8,H11-H12,H15-H16,H18-H19].

Z analizy danych z opisanej wyżej bazy *KnotProt* wynika, iż 91% białek o nietrywialnej topologii pełni funkcję enzymatyczną [H12]. Ponadto okazuje się, iż położenie węzła jest skorelowane z centrum aktywnym (tę własność wykazałam analizując szczegółowo trzy rodziny zawężonych białek). W połączeniu z przedstawionym wcześniej faktem, iż topologia białka jest zachowywana silniej niż sekwencja oraz struktura trzeciorzędowa, możemy łatwo wywnioskować, że na położenie centrum aktywnego białka oraz jego ścieżkę sygnałową (np.

wewnętrzny molekularny przekaz sygnału między białkiem, ligandem, oraz tRNA w procesie metylacji w białkach zawężlonych z rodziny SpoUT) istotny wpływ mają więzy topologiczne. Ilustracją tego stwierdzenia jest np. wynik przedstawiony w pracy [H12], w której pokazałam, iż konfiguracja slipknota pełni funkcję “paska” spinającego transmembranowe helisy – struktura ta stabilizuje utworzony helikalny kanał podczas aktywnego transportu jonów, który wymaga dużych zmian konformacyjnych.

Kolejną ważną rolę węzła można zidentyfikować badając korelacje pomiędzy występowaniem zawężeń w białkach a własnościami środowiska, w jakim one żyją. Zgodnie ze szczegółowymi danymi przedstawionymi na stronie *KnotProt* [H12], większość zawężlonych białek funkcjonuje w ekstremalnych warunkach, np. w gorących źródłach, czy też przy bardzo wysokim lub niskim pH. Takie własności sugerują, iż topologia pełni także rolę chroniącą białka, a w szczególności ich centra aktywne, przed termiczną lub chemiczną degradacją. Większą stabilność białka zawężlonego podczas termicznej oraz mechanicznej degradacji, w porównaniu do rozwężlonego białka o tym samym foldzie, strukturze, i bardzo dużym podobieństwie sekwencyjnym, wykazałam w pracy [H19]. W pracach [H15,H19] wykazałam też, iż podczas mechanicznego rozciągania węzeł niespodziewanie zaciska się w okolicach stanu natywnego, natomiast we współpracy z doświadczalnikami pokazałam, iż podobne zjawisko występuje także podczas chemicznej denaturacji [H8]. Istnienie węzłów w stanie zdenaturowanym zostało także zapostulowane w kilku kolejnych pracach doświadczalnych (25).

Z powyższych rozważań wynika, iż doświadczalne rozwężlenie węzła w białku – czyli jedno z istotnych wyzwań w tej dziedzinie dyskutowane od kilku lat – wydaje się niemożliwe bez uszkodzenia całej struktury białka. Jest to dosyć zaskakujący wniosek. Niemniej jednak ta unikalna cecha ma kluczowe znaczenie dla czasu życia białka. Z jednej strony wiemy, że prawdopodobieństwo spontanicznego zawężlenia lub rozwężlenia białka jest bardzo niskie. Z drugiej strony prawdopodobieństwo poprawnego zwinięcia białka ze stanu (wstępnie) zawężlonego jest niemalże stuprocentowe [H17]. Implikuje to, iż obecność węzła wydłuża czas życia białka w niekorzystnym środowisku, gdyż cząstkowe rozwinięcie struktury pod wpływem termicznej lub chemicznej denaturacji wyprowadza białko tylko na „krótką chwilę” ze stanu natywnego. Kontynuując tę myśl można zapostulować, iż trudności w procesie zwijania są rekompensowane poprzez wydłużenie okresu czasu w którym białko jest biologicznie aktywne.

Niezależnie od powyższych wniosków, w rozdziale V omówię wpływ nietrywialnej topologii na własności procesu mechanicznego rozwijania białek z węzłami oraz slipknotami.

III.D Rola lassa w białkach

Charakterystykę występowania i rolę lassa w białkach opisałam w pracach [H1,H4,H13]. W pracy [H1] na podstawie biologicznej klasyfikacji wykazałam, iż białka z lassami występują znacznie częściej w wirusach, roślinach i grzybach niż w innych królestwach. Ponadto, w przeciwieństwie do białek z węzłami i slipknotami, tylko 38% białek z lassami pełni funkcje enzymatyczne. Wykazałam, iż poszczególne funkcje białek są typowe dla konkretnego motywu lassa. Ponadto zidentyfikowałam przykłady białek, w których biologiczna funkcja może być wspierana przez motyw lassa; w szczególności:

- w anhidrazach, które posiadają bardzo płytki węzeł (termiczne fluktuacje są wystarczające aby rozwęzić białko), motyw lassa podtrzymuje stabilną konformację centrum aktywnego,
- w rodzinie białek z aktywnością RNase (1 RNase H z hipertermofilnych archeonów organizmów żyjących w temperaturze powyżej 100 °C) lasso typu L₃ jest prawdopodobnie odpowiedzialne za bardzo wysoką termiczną stabilność białka,

poprzez nałożenie więzów na nieustrukturyzowane centrum aktywne. Podobną rolę może pełnić motyw lassa w białkach nieustrukturyzowanych z topologią L_1 , które stanowią około 5% białek.

Te jak i inne przykłady są zaprezentowane w stworzonym przeze mnie serwerze i bazie danych *LassoProt*, dostępnym na stronie internetowej <http://LassoProt.cent.uw.edu.pl> [H1].

Znaczenie biologiczne zawężenia typu lasso potwierdziłam bezpośrednio w szczególności dla leptyny. Leptyna, która odgrywa kluczową rolę w regulacji pobierania pokarmu i gospodarki energetycznej organizmu, jest pierwszym białkiem, w którym zidentyfikowałam motyw lassa [H13]. Po związaniu leptyny z receptorami w podwzgórzu (część mózgu u wszystkich kręgowców), neurony wytwarzają neurotransmitter, który decyduje o wzbudzeniu lub wyhamowaniu apetytu (na podstawie *Nat Clin Pract Endocrinol Metab* **2** (6): 318–27). W pracy [H13] na podstawie badań doświadczalnych i teoretycznych pokazałam, iż mostek cysteinowy (odpowiedzialny za istnienie motywu lassa) odgrywa ważną rolę w wiązaniu receptora, a więc pośredniczy w aktywowaniu biologicznej funkcji poprzez tłumienie lokalnych drgań w okolicach wiązania receptora, chociaż miejsce to nie jest bezpośrednio związane z mostkiem cysteinowym. Ponadto na podstawie białek o tym samym zwoju i motywie topologicznym L_1 pokazałam, iż zmiany w dynamice stanu natywnego (w okolicach centrum aktywnego) są determinowane i zależą od położenia oraz rozmiaru pętli [H4].

Niezależnie od tych wyników w pracy [H1] pokazałam, iż motywy lassa istnieją w mini-białkach (27), które są obecnie powszechnie modyfikowane w celu uzyskania konkretnych farmakologicznych własności. Motyw lassa i opracowane przeze mnie metody do jego identyfikacji dostarczają nowych narzędzi do modyfikacji tych mini-białek oraz mogą prowadzić do nowych terapeutycznych rozwiązań nie tylko w przypadku białek z mostkiem cysteinowym.

Podsumowując, w pracach [H1,H4,H13] pokazałam, iż struktura lassa jest dodatkowym elementem, który kontroluje funkcję i krajobraz energetyczny białka (omówiony w paragrafie IV). Struktura ta powinna też być wzięta pod uwagę np. podczas projektowania leków, chociażby dla białek z foldem wiązki czterech helis (ang. „four helix bundle”) [H4], czy mini-białkach [H1]. Warto także podkreślić, iż w przeciwieństwie do białek zawężonych pojedyncza mutacja cysteiny powoduje zmianę topologii w białkach lassowych. Niemniej jednak właśnie ta własność może być wykorzystana do sterowania aktywnością biologiczną białka, poprzez zmianę topologii za pomocą pojedynczych mutacji lub zmieniając warunki z utleniających na redukcyjne.

IV Krajobraz energetyczny, powstawanie węzłów, slipknotów oraz lass

W tej części wykażę, iż proces zwijania białek o nietrywialnej topologii, w przeciwieństwie do białek o topologii trywialnej, przypomina reakcję chemiczną o jednoznacznej kolejności zdarzeń których zajście jest determinowane konformacją stanów przejściowych, a krajobraz energetyczny nie jest idealnie gładki, i dla zdecydowanej większości białek nadal jest niewyznaczalny obecnymi modelami teoretycznymi. Ponadto pokażę, iż konwencjonalne układy współrzędnych stosowane do opisu stanu przejściowego podczas zwijania nie mogą być zastosowane dla białek zawężonych. Do zbadania krajobrazu energetycznego białek zawężonych użyłam modeli gruboziarnistych o różnej ziarnistości bazujących na strukturze natywnej białka (tzw. Structure Based Models – SBM) oraz pełnoatomowego modelu z jawnym rozpuszczalnikiem (Desmond – Schrödinger).

Mechanizmy zwijania zawężonych białek omówię dla przypadku białek z głębokim węzłem

(w punkcie **A**), białek z płytkim węzłem, wraz ze szczegółową interpretacją ich krajobrazu energetycznego (w punkcie **B**), oraz białek z lassami (w punkcie **C**).

IV. A Mechanizmy zwijania białek z głębokim węzłem

Modele teoretyczne zwijania białek wykorzystujące informację o ich stanie natywnym działają dzięki założeniu, iż białka posiadają minimalnie sfrustrowany krajobraz energetyczny. Modele takie są obecnie jednym z podstawowych narzędzi służących do zrozumienia, jak relatywnie słabe molekularne oddziaływania prowadzą do szybkiego i kooperatywnego zwijania białek (11,12). Ponieważ zarówno mechanizmy zwijania białek, jak też ich ruchy funkcyjne wynikają z tego samego krajobrazu energetycznego, oznacza to, że analiza tego krajobrazu powinna pozwolić na scharakteryzowanie relacji pomiędzy strukturą, procesem zwijania, a funkcją białka. Badania teoretyczne oparte na takich założeniach, prowadzone w ostatnich latach, ujawniły wiele nowych mechanizmów odpowiedzialnych za regulację biologicznej funkcji pojedynczych białek (a także maszyn molekularnych, np. rybozomu), jednocześnie pozwalając na uniknięcie znacznej części kosztownych oraz trudnych w interpretacji badań doświadczalnych (12). Okazuje się jednak, że w przypadku białek zapętionych należy opisać tu podejście zmodyfikować.

W standardowym ujęciu teoria lejkowatego profilu energii swobodnej podczas zwijania zakłada istnienie zbioru dróg, którymi biomolekuła może osiągnąć natywną konformację. Liczba możliwych konformacji zmniejsza się wraz ze zbliżaniem do stanu natywnego białka, przy czym siłą sterującą formowanie się struktury są natywne kontakty. Zwijanie zaczyna się od stworzenia natywnych drugorzędowych elementów struktury białka (centrum zwijania), które jest stopniowo lokalnie rozbudowywane poprzez proces kondensacji, aż do momentu utworzenia trzeciorzędowej natywnej struktury białka. Zwijanie postępuje po relatywnie gładkim krajobrazie energetycznym. Stwierdzenia te nie są jednak zgodne z tym, jak zachowują się białka zapętione, i w związku z tym analiza procesu ich zwijania wymaga modyfikacji wspomnianej tu teorii. W moich pracach zapostulowałam, iż taka modyfikacja wymaga uwzględnienia w krajobrazie energetycznym zapętionych białek barier topologicznych, których efektywne pokonywanie wymaga specyficznych oddziaływań w białkach z głębokim węzłem. Jak opiszę poniżej, w wyniku moich badań zidentyfikowałam takie bariery oraz opisałam ich własności.

Warto także wspomnieć o podejściu eksperymentalnym, w którym w początkowej fazie rozwoju tej dziedziny niepoprawnie założono, iż białka z węzłem zwijają się i rozwijają bez większych problemów. Z takim właśnie stanem wiedzy zaczęłam badać mechanizmy zwijania białek zawężonych. Dziś jednak w dalszym ciągu nie umiemy rozwęzić białek doświadczalnie (a zatem okazało się to bardzo problematyczne), ani nie dysponujemy eksperymentalnymi narzędziami do określenia ich topologii podczas zwijania lub rozwijania. W związku z tym podejście teoretyczne jest jedynym dostępnym narzędziem do badania relacji pomiędzy zapętlaniem a konformacją białka, i to dzięki jego zastosowaniu udało się poznać mechanizmy zwijania zapętionych białek.

Najważniejsze wyniki moich badań dotyczących procesu zwijania białek z głębokim węzłem są następujące. W pracach [H10,H14,H16,H17] wykazałam, iż białka zawężone – z trzech grup o różnym foldzie i funkcji – mimo wcześniejszych założeń (28) mogą się same zawęzić w modelach typu SBM. W pracy [H17] pokazałam, iż natywne kontakty są wystarczające aby białka (oznaczane YibK i YbeA, z rodziny SPOUT) same się zawęziły, jednakże prawdopodobieństwo ich zawężenia jest bardzo niskie, a krajobraz lejka zwijania jest niespodziewanie wąski (posiada tylko jedną wiązkę ścieżek), czyli jest bardzo odmienny od krajobrazu trywialnych globularnych białek. Proces zwijania wspomnianych tu białek składa się z pięciu etapów, z których trzy pierwsze to zmiany konformacyjne odpowiadające I i II ruchowi Reidemeistera (dobrze znanym w teorii węzłów). Zapętlenie białka następuje w wyniku powstania (w czwartym etapie) natywnie skreconej pętli,

przez którą (w ostatnim, piątym etapie) przeciągany jest krótszy koniec białka w nienatywnej konformacji. Analiza czasów zwijania pokazuje, iż najczęściej czasu zajmuje oraz najmniej prawdopodobny jest ostatni etap, czyli przeciągnięcie końca białka przez pętlę, odpowiadający za zmianę topologii. Właśnie to przejście nazwałam barierą topologiczną.

W przeciwieństwie do węzłów powstających spontanicznie i w przypadkowych miejscach w polimerach, węzły w białkach powstają zawsze w tym samym miejscu, natomiast ich utworzenie jest krokiem hamującym proces zwijania. Mimo że wyznaczone przeze mnie prawdopodobieństwo zawężenia jest bardzo niskie, to wyniki te, otrzymane w roku 2009 w pracy [H17], były pewnego rodzaju przełomem w tej dziedzinie, gdyż wbrew interpretacji wcześniejszych wyników doświadczalnych pokazały, iż czasy zwijania tych białek są zaskakujące długie i są determinowane momentem zapętlenia. W roku 2012 moje badania teoretyczne zostały potwierdzone eksperymentalnie przez prof. S. Jackson (Cambridge University), która pokazała, iż czas samozawężenia wynosi około 20 minut, a nie kilka sekund. Jak się okazuje, kilkusekundowe czasy zawężania białek obserwowane eksperymentalnie mogą być osiągnięte tylko w przypadku procesu zwijania zaczynającego się z konfiguracji zawierającej nierozwężlony stan zdenaturowany, co także wykazałam w pracy [H17].

W pracy [H11] pokazałam, iż skonstruowane laboratoryjnie (poprzez duplikację domen) zawężone białko, w przeciwieństwie do białek z rodziny SpoUT o podobnej długości, zwija (zawęzła się) i rozwija się w modelu SBM pomijając topologiczne pułapki, mimo że schemat jego zawężania pozostaje ten sam. Analizując teoretycznie możliwe mutacje (zmieniające własności poszczególnych aminokwasów oraz wielkość topologicznej bariery poprzez wydłużenie nieustrukturyzowanego elementu łańcucha peptydowego w pętli węzłowej) w rejonie bariery topologicznej pokazałam, w jaki sposób ewolucja mogłaby optymalizować proces zwijania. Wykazałam, iż białko z węzłem zwija się dziesięć razy wolniej niż odwężlone (dimer połączony mostkiem cysteinowym), zgodnie z doświadczalnymi przypuszczeniami. Wykazałam też, iż obserwowane doświadczalnie zakrzywienie w krzywych zwijania oraz rozwijania w funkcji termicznej lub chemicznej denaturacji (tzn. na wykresie chevron, przedstawiającym czas zwijania/rozwijania w zależności od temperatury) jest poprawnie odtwarzane w modelu SBM, oraz że zawężenie jest ograniczone do wąskiego zakresu temperatur w porównaniu do odwężlonego białka. Na podstawie obszernej analizy czasów zwijania można sądzić, iż zakrzywienie na wykresie chevron jest konsekwencją przejścia z niedyfuzyjnej, powolnej dynamiki do dyfuzyjnej. Wyniki tej pracy zostały w 2015 roku potwierdzone przez szczegółową analizę doświadczalną wykonaną przez grupę prof. S. Jackson.

W pracy [H14] pokazałam, iż krajobraz energetyczny najmniejszego białka z węzłem jest także minimalnie sfrustrowany. Warto zauważyć, iż historycznie było to pierwsze białko dla którego własności termodynamiczne, wraz ze zmianami topologicznymi, zostały wyznaczone oraz opracowane metodami do charakterystyki krajobrazu energii swobodnej (The Weighted Histogram Analysis Method, WHAM). W pracy tej pokazałam, iż niezależnie od ziarnistości użytego modelu (tzn. reprezentacji aminokwasów albo przez sam węgiel $C\alpha$, albo też wszystkie ciężkie atomy) węzeł powstaje zawsze w ostatnim etapie zwijania, po pokonaniu topologicznej bariery. Ponadto w pracy tej pokazałam, iż płytkie węzły w stanie zdenaturowanym są bardzo rzadkie, a głębokie nie są obserwowane. Wykazałam także, iż prawdopodobieństwo zawężenia, wybór końca białka oraz geometrii (trywialnej lub slipknotowej) w jakiej topologiczna bariera zostanie pokonana, są determinowane głębokością węzła. Analizując proces zwijania dla białka o takim samym zwoju ale o różnej głębokości węzła, w pracy [H5] wykazałam słuszność tej tezy, przynajmniej w modelach typu SBM.

Warto podkreślić, iż w celu zweryfikowania otrzymanych wyników modelami SBM podjęłam próbę zwinięcia najmniejszego białka w modelu CABS, stworzonym przez prof. Kolińskiego, którego hamiltonian jest zbudowany na podstawie potencjałów statystycznych (i symuluje zwijanie białek techniką de novo). Badania te pokazały, iż prawdopodobieństwo osiągnięcia stanu

zawężłonego w tym modelu wynosi 20%, podczas gdy w opracowanym przeze mnie modelu SBM wynosi ono ponad 90%.

Podsumowując, ważnym wnioskiem wynikającym z przeprowadzonych przeze mnie badań jest charakterystyka procesu zwijania białek, z uwzględnieniem procesu ich zapętlenia. Z analizy procesów kinetycznych i termodynamicznych zawężłonych białek wynika, iż białka zawężłają się tworząc natywną pętlę przez którą jest przeciągany krótszy koniec węzła, w zależności od jego długości w konformacji zgiętej (slipknot) lub trywialnej. Ten niesamowicie prosty – jednakże nieintuicyjny z punktu widzenia spontanicznego i przypadkowego zawężłania polimerów – proces jest łatwy do wytłumaczenia, jeśli zauważy się, że występujące w białkach węzły są typu skręconego („twist knots”) [H12]. Proces ten był już zapostulowany w [10,31], choć w tamtym czasie znacznie mniej białek z węzłami było znanych, w szczególności nie było jeszcze odkryte białko z węzłem 6_1 . W węzłach skręconych redukcja jednego skrzyżowania powoduje całkowite rozwężlenie, a więc węzły 3_1 , 4_1 , 5_2 , 6_1 występujące w białkach mogą powstać poprzez stworzenie odpowiednio skręconej pętli, a następnie jednorazowe przeciągnięcie przez nią jednego z końców białka. W pracy [H16] wykazałam słuszność tej hipotezy, analizując proces zwijania białka ze skomplikowanym węzłem 6_1 , choć w tym wypadku prawdopodobieństwo poprawnego zwinienia jest tutaj jeszcze niższe niż dla białka z rodziny SpoUT. Warto również wspomnieć, iż ostatnie eksperymenty prof. S. Jackson (Cambridge University) pokazują, iż czas zwijania białek z rodziny SpoUT może być zredukowany aż o rząd wielkości przez białka opiekuńcze (czaperony) – może to wyjaśniać, czemu w prowadzonych przeze mnie analizach teoretycznych, nie uwzględniających obecności takich białek opiekuńczych, prawdopodobieństwa zawężłania są bardzo małe. Moje kolejne prace, które już nie wchodzi w skład tego dzieła, pokazują, iż także w czaperonie powinien zachodzić ten sam mechanizm zwijania. Kolejnym argumentem przemawiającym za opisanym tu mechanizmem zawężłania białek jest fakt niezidentyfikowania do tej pory białek z węzłem typu 5_1 – węzeł ten nie jest typu skręconego i musiałby powstać po dwukrotnym przewleczeniu przez skręconą pętlę jednego z końców. Potwierdza to zatem tezę, iż w procesie zwijania węzeł powstaje w wyniku jednorazowego przeplecenia jednego z końców białka przez uformowaną wcześniej natywną pętlę.

IV. B Szczegółowa analiza zwijania najmniejszego białka z węzłem

Modele gruboziarniste typu SBM umożliwiają zbadanie własności termodynamicznych biomolekuły w przypadku gładkiego krajobrazu energetycznego. Jednakże, jak pokazałam powyżej, moment pokonywania topologicznej bariery jest związany z ciasnym nienatywnym upakowaniem struktury i pojawieniem się wielu nienatywnych kontaktów. Te nienatywne kontakty są uwzględniane w modelu SBM jedynie poprzez efekt wyłączonej objętości, prawdopodobnie spowalniając lub blokując proces zwijania przez zwiększenie nierówności na powierzchni energii. W pracy [H9] na podstawie symulacji z pełnoatomową reprezentacją białka z jawnym rozpuszczalnikiem (przeprowadzonych na superkomputerze Anton, D. Shaw programem Desmond) po pierwsze wykazałam, iż proces spontanicznego samozawężłania przez konformację slipknota jest możliwy. Po drugie pokazałam, iż te same oddziaływania są siłą napędową podczas zawężłania w obu modelach. Analiza mechanizmu zwijania pokazała, iż nienatywne kontakty, w przeciwieństwie do założenia, iż pełnią rolę tylko naprowadzającą przeciągany koniec na skręconą pętlę (28), uczestniczą aktywnie w przeprowadzaniu elementu białka przez jego natywną pętlę. Zestawienie wyników prac [H9] i [H14] pokazuje, iż model SBM może być użyty do analizy białek z węzłami, przynajmniej tych o podobnym foldzie.

Niemniej jednak w pracy [H2] pokazałam (używając twierdzenia Bayesa, konstruując prawdopodobieństwo warunkowe pomiędzy stanami równowagowymi i konformacją białka w okolicach maksimum na krajobrazie energetycznym $F(Q)$), iż liczba natywnych kontaktów Q , powszechnie używana do analizy białek o trywialnej topologii (29), nie opisuje poprawnie procesu

przechodzenia przez topologiczną barierę (ang. „transition state”), a określa jedynie postęp reakcji. Ponadto zastosowane przeze mnie metody wariacyjne do optymalizacji wyboru układu współrzędnych także nie pozwalają na wyznaczenie konformacji białka odpowiadającej stanowi przejściowemu, w przeciwieństwie do białek globularnych (18). Wynik ten jest zaskakujący, gdyż coraz większa liczba prac pokazuje, iż natywne kontakty determinują zwijanie białek także w pełnoatomowych symulacjach (18). Reasumując, zestawienie moich prac pokazuje, iż mimo że natywne kontakty są wystarczające do pokonania topologicznej bariery oraz proces zawężania przebiega tak samo w modelu SBM jak i z jawnym rozpuszczalnikiem, to liczba natywnych kontaktów nie jest dobrym układem współrzędnych do opisu tego zjawiska. Są to istotne wnioski, które wskazują możliwe kierunki dalszych badań.

Kolejnym zasadniczym elementem, który decyduje o gładkości krajobrazu energetycznego w modelach SBM, jest liczba oraz sposób doboru natywnych kontaktów. W pracach [H2,H5] scharakteryzowałam minimalną liczbą fizycznych kontaktów (kilka reprezentacji, uwzględniając także oddziaływania nienatywne bazując na koewolucji par aminokwasów, Direct Coupling Analysis) potrzebną do przeprowadzenia białka przez topologiczną barierę i jego zwinięcie do stanu natywnego. Możliwość skonstruowania minimalnie sfrustrowanej mapy (krajobrazu) oddziaływań dla najmniejszego zawężonego białka sugeruje, iż podobna mapa może istnieć dla białek z głębokim węzłem (np. białek z rodziny SpoUT omówionych powyżej), jednakże obecnie nie potrafimy jej wyznaczyć. Minimalnie sfrustrowana sekwencja (o gładkim krajobrazie energetycznym) także w przypadku białek z głębokim węzłem może wynikać z procesu ewolucji (selekcji specyficznej sekwencji), który zapewniłby białkom węzłowym efektywny proces zwijania, podczas gdy proces przyspieszonej denaturacji wynikający z relatywnie małej liczby stabilizujących natywnych oddziaływań byłby chroniony przez więzy topologiczne.

IV. C Krajobraz energetyczny białek z lassami

Białka z lassami stanowią kolejną grupą zapętlonych biomolekuł. Analiza tych białek w stanie utlenionym i zredukowanym (odpowiadających odpowiednio obecności lub brakowi lassa) jest ważnym narzędziem, które pozwala zrozumieć również pewne własności białek z węzłami.

W pracach [H4,H13], analizując białka o tym samym foldzie zawierające (w pięciu przypadkach) pojedyncze lasso typu L_1 , lub (w trzech przypadkach) nie zawierające motywu lassa, wykazałam, że krajobraz energetyczny białek lassowych jest minimalnie sfrustrowany. Natywne kontakty okazują się wystarczające do pokonania topologicznej bariery, która w przypadku lassa odpowiada przeciągnięciu końca białka przez zamkniętą kowalencyjną pętlę. Na podstawie opracowanych przeze mnie narzędzi do analizy geometrii lass pokazałam, iż podczas pokonywania bariery topologicznej, wraz ze wzrostem długości zamkniętej pętli, konformacja przeciąganego końca zmienia się z zagiętej (przypominającej slipknot) do wyprostowanej (ang. „plug”). Pokonanie topologicznej bariery jest elementem determinującym czas oraz prawdopodobieństwo zwinięcia.

Szeroka analiza (przy użyciu modelu SBM w reprezentacji węgli $C\alpha$ oraz pełnoatomowej) pojedynczego białka (leptyny) pokazała [H4], iż krajobrazy energetyczne oraz ścieżki zwijania dla białka w stanie zredukowanym oraz utlenionym są zaskakująco do siebie podobne. Z kolei porównanie wyników doświadczalnych i teoretycznych pokazało niedoszacowanie proporcji pomiędzy wielkością bariery energii swobodnej w funkcji liczby kontaktów dla trzech układów: zredukowanego bez możliwości tworzenia mostka w procesie zwijania/rozwijania, zredukowanego z możliwością utworzenia mostka, oraz utlenionego (dla którego mostek istnieje podczas całego procesu symulacji). Różnica ta prawdopodobnie wynika z faktu, iż liczba natywnych kontaktów Q nie może być użyta jako współrzędna reakcji do opisu krajobrazu energetycznego białek o nietrywialnej topologii. Jest to zgodne z wynikami pracy [H2], w której na podstawie analizy zwijania zawężonego białka wykazałam, iż liczba kontaktów Q rzeczywiście nie może być wykorzystana do wyznaczenia stanu przejściowego.

Podsumowując, otrzymane wyniki pokazują, iż także w przypadku białek lassowych ewolucja zaprojektowała minimalnie sfrustrowaną sekwencję, a w konsekwencji gładki krajobraz energetyczny. Ponadto proces zwijania jest złożonym procesem, który może być badany modelem SBM, jednakże poprawny opis stanu przejściowego wymaga użycia współrzędnych innych niż liczba natywnych kontaktów Q czy też średnie kwadratowe odchylenie (RMSD). Rolę takich współrzędnych może odgrywać rozważany w pracy [H1] rozmiar powierzchni minimalnej.

V Krajobraz energetyczny – rozwijanie i rozplątywanie węzłów oraz slipknotów

W powyższych rozdziałach przedstawiłam własności krajobrazu energetycznego z punktu widzenia kinetyki i termodynamiki zwijania białek. W obecnym rozdziale omówię własności krajobrazu energetycznego z perspektywy rozwijania i konformacji białka w warunkach zdenaturowanych. Na podstawie symulacji komputerowych i danych doświadczalnych wykażę [H8], iż rozsoplanie białka jest skomplikowanym procesem (co potwierdza fakt, iż wciąż nie został on zrealizowany doświadczalnie). Niemniej jednak, na podstawie teoretycznej analizy procesu mechanicznego rozciągania przedstawię model rozsoplania węzła w białku [H15]. Ponadto wykażę istnienie quasi-stabilnego stanu w białku ze slipknotem [H18].

V. A Zawężone białka w stanie zdenaturowanym oraz model ich rozwiązywania

Termiczna oraz chemiczna denaturacja białka są bardzo ważnymi procesami, wiążącymi unikalną aktywną biologicznie konformację białka ze stanem kłęбка statystycznego. Badanie tych procesów wymaga spełnienia warunku odwracalności denaturacji. W przypadku białek o trywialnej topologii warunek ten może nie być spełniony jeśli wystąpi zjawisko agregacji.

W pracy [H12] na podstawie szerokich badań doświadczalnych i teoretycznych wykazałam, iż rozwężenie białka jest procesem blokującym odwracalną denaturację. Analiza danych kinetycznych otrzymanych z symulacji komputerowych, w których topologia może być jednoznacznie określona, pokazała, iż rozwijanie jest procesem co najmniej trzyetapowym. Ostatnim etapem takiego procesu zawsze jest zaciśnięcie się węzła w okolicach jego natywnego położenia. Czas rozsoplania jest co najmniej o rząd wielkości wolniejszy niż rozwijania. Zaobserwowana doświadczalnie podczas chemicznej denaturacji pętla histerezy wynika z istnienia dwóch możliwych procesów zwijania: albo ze stanu zdenaturowanego ale zawężonego, albo też (choć bardzo rzadko) ze stanu zdenaturowanego i rozwężonego. Wyniki te jednoznacznie pokazują, iż (niestety) unikalną cechą białek zawężonych (z głębokim węzłem) jest zachowanie węzła w stanie zdenaturowanym. Wniosek ten został ostatnio pośrednio potwierdzony przez niezależne badania eksperymentalne przeprowadzone przez prof. S. Jackson (Cambridge University).

Przedmiotem pracy [H15] było skonstruowanie teoretycznego modelu mechanicznego rozciągania, określającego prawdopodobieństwo rozsoplania białka z rodziny SpoUT. Podstawową zaletą metody mechanicznego rozciągania jest obniżenie bariery energii swobodnej (przy spełnieniu założenia, iż kierunek rozciągania jest zgodny z naturalnym, termodynamicznym układem współrzędnych), a w konsekwencji znacznie skuteczniejsze przeszukiwanie krajobrazu energetycznego i wyznaczenie statystyki badanego zjawiska, która pozwala na wyznaczenie wszystkich stanów przejściowych. W pracy [H15] na podstawie analizy wyników symulacji – w modelu SBM – rozciągania białka z węzłem w różnych kierunkach, poprzez wybór różnych par aminokwasów za które białko jest rozciągane, w szerokim zakresie temperatur oraz prędkości rozciągania, wyznaczyłam prawdopodobieństwo rozwężenia białka. Ponadto zaproponowałam ogólny wzór określający prawdopodobieństwo rozsoplania białek z rodziny SpoUT w wyniku mechanicznego rozciągania. Opracowany model ma kluczowe znaczenie dla dalszych prac

doświadczalnych przy użyciu metod AFM i szczypców optycznych, które powinny doprowadzić do opracowania skutecznej techniki rozwiązywania węzłów w białkach.

V. B Unikalne metastabilne własności białek slipknotowych

Z matematycznego punktu widzenia slipknot ma trywialną topologię i wydawać by się mogło, iż pociągnięcie za jego końce spowoduje rozciągnięcie całej struktury do postaci prostego odcinka. W pracy [H18] modelując teoretycznie (w modelu SBM) proces rozciągania białka ze slipknotem z różnymi siłami pokazałam, iż oprócz pełnego rozciągnięcia slipknota możliwe jest także osiągnięcie konfiguracji, w której slipknot jest zakleszczony. W szczególności pokazałam, w jaki sposób osiągnięcie jednej z tych konfiguracji (rozciągniętej lub zakleszczonej) determinowane jest przez parametry rozciągania, takie jak jego kierunek oraz miejsce zaczepienia, wartości siły rozciągającej oraz prędkości ciągnięcia. Wykazałam też, iż w przypadku dużej prędkości lub siły rozciągania możliwy jest dwuetapowy proces prowadzący do pełnego rozciągnięcia białka, w którym pojawia się stan pośredni (metastabilny) z zakleszczonym slipknotem, charakteryzujący się dużą siłą oporu lub długim okresem czasu oczekiwania na przejście do kolejnego stanu. Taki dwuetapowy proces rozwijania może być opisany równaniem Bella. Ponadto na podstawie analizy energii wygięcia, giętkości łańcucha, oraz siły tarcia łańcucha peptydowego w slipknotowej i węzłowej pętli skonstruowałam teoretyczny model opisujący warunki konieczne do pojawienia się stanu metastabilnego. Należy podkreślić, iż powyższe wyniki zostały potwierdzone doświadczalnie (30) przez grupę prof. Hongbin Li (University of British Columbia, Canada). Mimo że ten teoretyczny model został sformułowany dla białek globularnych, stanowi on także cenne narzędzie badań nad białkami membranowymi, których znaczna część posiada konformację slipknota.

Literatura

- [1] M.L. Mansfield, *Are there knots in proteins?*, Nat Struct Biol (1994) 1:213–214.
- [2] W.R. Taylor, *A deeply knotted protein structure and how it might fold*, Nature (2000) 406:916–919.
- [3] R.C. Lua, A.Y. Grosberg, *Statistics of knots, geometry of conformations, and evolution of proteins*, PLoS Comput Biol (2006) 2:e45.
- [4] P. Virnau, L.A. Mirny, M. Kardar, *Intricate knots in proteins: Function and evolution*, PLoS Comput Biol (2006) 2:e122.
- [5] N.P. King, E.O. Yeates, T.O. Yeates, *Identification of rare slipknots in proteins and their implications for stability and folding*, J Mol Biol (2007) 373:153–166.
- [6] K. Shimokawa, K. Ishihara, I. Grainge, D. Sherratt, M. Vazquez, *FtsK-dependent XerCD-dif recombination unlinks replication catenanes in a stepwise manner*, PNAS (2013) 110, 52: 20906-20911.
- [7] P.Y. Chou, G.D. Fasman, *Prediction of protein conformation*, Biochemistry (1974) 13, 2: 222–245.
- [8] K.L. Tkaczuk, S. Dunin-Horkawicz, E. Purta, J.M Bujnicki, *Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases*, BMC Bioinformatics (2007) 8:73.
- [9] R. Potestio, C. Micheletti, H. Orland, *Knotted vs. Unknotted Proteins: Evidence of Knot- Promoting Loops*, PloS Comput Biol (2010) 6:e1000864.
- [10] J. I Sulkowska, *Stretching and folding proteins in coarse grained models*, doktorat (2008)
- [11] P.G. Wolynes, *Evolution, energy landscapes and the paradoxes of protein folding*, Biochimie (2014) S0300-9084, 14:00389-7.
- [12] P.C. Whitford, J.N. Onuchic, *What protein folding teaches us about biological function and molecular machines*, Current Opinion Struct Bio (2015) 30, 57–62.
- [13] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, *Funnels, pathways, and the energy landscape of protein folding: a synthesis*, Proteins (1995) 21, 3:167-95.
- [14] G.D. Rose, P.J. Fleming, J.R. Banavar, A. Maritan, *A backbone-based theory of protein folding*, PNAS (2006) 103, 45:16623–16633.
- [15] K.A. Dill, *Dominant forces in protein folding*, Biochemistry (1990) 29 (31): 7133–7155.

- [16] A. Godzik, A. Kolinski, J. Skolnick, *Topology fingerprint approach to the inverse protein folding problem*, J Mol Biol (1992) 227, 1, 227-238.
- [17] C. Levinthal, *Are there pathways for protein folding?*, Journal de Chimie Physique et de Physico-Chimie Biologique (1968) 65: 44-45.
- [18] R.B. Best, G. Hummer, W.A. Eaton, *Native contacts determine protein folding mechanisms in atomistic simulations*, PNAS (2013) 110(44): 17874-17879.
- [19] E.R. Henry, R.B. Best, W. A. Eaton, *Comparing a simple theoretical model for protein folding with all-atom molecular dynamics simulations*, PNAS (2013) 110, 44, 17880-17885.
- [20] K.C. Neuman, A. Nagy, *Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy*, Nat Methods (2008) 5(6): 491-505.
- [21] J.M. Fernandez J.M, H. Li, *Force-clamp spectroscopy monitors the folding trajectory of a single protein*, Science (2004) 12;303(5664):1674-8.
- [22] H. Lu, B. Isralewitz, A. Krammer, V. Vogel, K. Schulten, *Unfolding of Titin Immunoglobulin Domains by Steered Molecular Dynamics Simulation*, Biophys. J. (1998) 75, 2, 662-671.
- [23] J.I. Sułkowska, M. Cieplak, *Stretching to understand proteins - a survey of the protein data bank*, Biophys J. (2008) 94,1: 6-13.
- [24] J.I. Sułkowska, F. Marcos, T. Hwa, J. N. Onuchic, *Genomics Aided Structure Prediction (GASP)*, PNAS (2012) 109, 26:10340-5.
- [25] A.L. Mallam, J. M. Rogers, S. E. Jackson, *Experimental detection of knotted conformations in denatured proteins*, PNAS (2010) 107: 8189-8194.
- [26] N.P. King, A.W. Jacobitz, M.R. Sawaya, L. Goldschmidt, T.O. Yeates, *Structure and folding of a designed knotted protein*, PNAS (2010) 107: 20732-20737.
- [27] Yanyan Li, Séverine Zirah, Sylvie Rebuffat, *Lasso Peptides: Bacterial Strategies to Make and Maintain Bioactive Entangled Scaffolds*, (2014) Springer, Oct 21
- [28] S. Wallin, K.B. Zeldovich, E.I. Shakhnovich, *The folding mechanics of a knotted protein*, J Mol Biol (2007) 368, 3: 884-93.
- [29] R.B. Best, G. Hummer, *Reaction coordinates and rates from transition paths*, PNAS (2005), 102, 19: 6732-7.
- [30] Ch. He, G. Lamour, A. Xiao, J Gsponer, H Li, *Mechanically Tightening a Protein Slipknot into a Trefoil Knot*, JACS (2014) 136 (34), 11946-11955
- [31] W.R. Taylor, *Protein knots and fold complexity: Some new twists*, Comput Biol Chem (2007) 31, 151-162

5. Omówienie pozostałych osiągnięć naukowo-badawczych

A) Granty i projekty badawcze

Granty badawcze – 7 jako kierownik, 2 jako główny wykonawca, 5 jako wykonawca
Granty na organizację konferencji – 7 jako kierownik / główny organizator, 4 jako partner

14. Rodzaj projektu: **Ideas Plus** (2 615 024 PLN)

Tytuł projektu: **Białka splecione – studium nowych struktur i rozwiązanie ich zagadki**

Rok rozpoczęcia realizacji: **2016**

Miejsce realizacji: **Wydział Chemii i Centrum Nowych Technologii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **Ministerstwo Nauki i Szkolnictwa Wyższego**

Charakter udziału habilitanta w projekcie: **kierownik projektu**

13. Rodzaj projektu: **Early Science** (czas obliczeniowy na superkomputerze Okeanos)

Tytuł projektu: **Wpływ wiązania substratu na dynamikę białka z węzłem**

Rok rozpoczęcia realizacji: **2016**

Miejsce realizacji: **Wydział Chemii i Centrum Nowych Technologii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **Interdisciplinary Centre for Mathematical and Computational Modeling (ICM), UW**

Charakter udziału habilitanta w projekcie: **kierownik projektu**

12. Rodzaj projektu: **EMBO Small Grant** (8400 Euro)

Tytuł projektu: **Synteza inhibitorów dla białek z rodziny TrmD**

Rok rozpoczęcia realizacji i zakończenia realizacji: **2016**

Miejsce realizacji: **Centrum Nowych Technologii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **European Molecular Biology Organization (EMBO)**

Charakter udziału habilitanta w projekcie: **kierownik projektu**

11. Rodzaj projektu: **EMBO Installation Grant** (250 000 Euro)

Tytuł projektu: **Comprehensive analysis of knotted proteins**

Rok rozpoczęcia realizacji: **2014**

Miejsce realizacji: **Centrum Nowych Technologii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **European Molecular Biology Organization (EMBO)**

Charakter udziału habilitanta w projekcie: **kierownik projektu**

10. Rodzaj projektu: **Projekt interdyscyplinarny „Skills/Inter”** (100 000 PLN)

Tytuł projektu: **Białko z węzłem jako węzeł gordyjski**

Rok rozpoczęcia realizacji i zakończenia realizacji : **2013 – 2014**

Miejsce realizacji: **Wydział Chemii i Centrum Nowych Technologii, Uniwersytet Warszawski (EMBO)**

Organ przyznający fundusze na realizację projektu: **Fundacja na rzecz Nauki Polskiej**

Charakter udziału habilitanta w projekcie: **kierownik projektu**

9. Rodzaj projektu: **Sonata Bis** (1 490 000 PLN)

Tytuł projektu: **Influence of knotted structure on function of proteins and protein structure prediction**

Rok rozpoczęcia realizacji: **2013**

Miejsce realizacji: **Wydział Chemii i Centrum Nowych Technologii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **Narodowe Centrum Nauki**

Charakter udziału habilitanta w projekcie: **kierownik projektu**

8. Rodzaj projektu: **Grant Homing-Plus** (253 000 PLN)

Tytuł projektu: **Entanglement in biology – how nature controls the topology of proteins**

Rok rozpoczęcia i zakończenia realizacji: **2012 – 2014**

Miejsce realizacji: **Wydział Chemii, Uniwersytet Warszawski we współpracy z University of California**

San Diego

Organ przyznający fundusze na realizację projektu: **Fundacja na rzecz Nauki Polskiej**

Charakter udziału habilitanta w projekcie: **kierownik projektu**

7. Rodzaj projektu: **Anton Grant PSCA00062P** (czas obliczeniowy – D. E. Shaw Research)

Tytuł projektu: **To knot or not to knot: slipknotting in the smallest knotted protein**

Rok rozpoczęcia i zakończenia realizacji: **2011-2012**

Miejsce realizacji: **University of California San Diego**

Organ przyznający fundusze na realizację projektu: **National Research Council (NRC) at the National Academy of Sciences of USA**

Charakter udziału habilitanta w projekcie: **główny wykonawca**

6. Rodzaj projektu: **Grant promotorski, nr N202 021 31/0739**

Tytuł projektu: **Zwijanie i rozwijanie białek w modelach gruboziarnistych**

Rok rozpoczęcia i zakończenia realizacji: **2006 – 2007**

Miejsce realizacji: **Instytut Fizyki Polska Akademia Nauki**

Organ przyznający fundusze na realizację projektu: **Ministerstwo Nauki i Szkolnictwa Wyższego**

Charakter udziału habilitanta w projekcie: **główny wykonawca**

Udział w grantach:

5. Rodzaj projektu: **Międzynarodowy, interdyscyplinarny, PHY-1212312 – prof. P Jennings**

Tytuł projektu: **Entanglement in Biology – How Nature Controls the Topology of Proteins**

Rok rozpoczęcia i zakończenia realizacji: **2012 – 2015**

Miejsce realizacji: **University of California San Diego, USA**

Organ przyznający fundusze na realizację projektu: **National Science Foundation of USA**

Charakter udziału habilitanta w projekcie: **wykonawca**

4. Rodzaj projektu: **Międzynarodowy, interdyscyplinarny – prof. E Radwan**

Tytuł projektu: **Theory and simulations of knotting in physical and biological systems ranging from proteins to glueballs**

Rok rozpoczęcia i zakończenia realizacji: **2011 – 2014**

Miejsce realizacji: **University of St. Thomas, USA**

Organ przyznający fundusze na realizację projektu: **National Science Foundation of USA**

Charakter udziału habilitanta w projekcie: **wykonawca**

3. Rodzaj projektu: **Międzynarodowy, interdyscyplinarny „Skills/Inter” – mgr W Niemyskiej**

Tytuł projektu: **Bańki mydlane – Czy biologię da się schwytać na matematyczne lasso?**

Rok rozpoczęcia i zakończenia realizacji: **2015 – 2016**

Miejsce realizacji: **Centrum Nowych Technologii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **Fundacja na rzecz Nauki Polskiej**

Charakter udziału habilitanta w projekcie: **wykonawca**

2. Rodzaj projektu: **Międzynarodowy, interdyscyplinarny, PHY- 1308264 – prof. J Onuchic**

Tytuł projektu: **Sharing the Energy Landscape for Folding an Function: from Proteins to Biomolecular Machines**

Rok rozpoczęcia i zakończenia realizacji: **2011 – 2016**

Miejsce realizacji: **University of California San Diego, Rice University (USA)**

Organ przyznający fundusze na realizację projektu: **National Science Foundation of USA**

Charakter udziału habilitanta w projekcie: **wykonawca**

1. Rodzaj projektu: **Międzynarodowy, interdyscyplinarny „Skills/Inter” – dr. hab. P Sułkowski**

Tytuł projektu: **On topology, interacting RNA, and quantum physics**

Rok rozpoczęcia i zakończenia realizacji: **2012 – 2013**

Miejsce realizacji: **Wydział Fizyki, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **Fundacja na rzecz Nauki Polskiej**

Charakter udziału habilitanta w projekcie: **wykonawca**

Granty na organizację konferencji – 7 jako kierownik / główny organizator, 4 jako partner

11. Rodzaj projektu: **Grant konferencyjny**

Tytuł projektu: **The Geometry and Topology of Knotting and Entanglement in Proteins**

Czas realizacji: **2017**

Miejsce realizacji: **Oaxaca, Meksyk**

Organ przyznający fundusze na realizację projektu: **Banff International Research Station (BIRS), Banff, Kanada**

Charakter udziału habilitanta w projekcie: **partner**

10. Rodzaj projektu: **Networking Grant, EMBO (5 000 Euro)**

Tytuł projektu: **VSSSB2016 sympozjum**

Czas realizacji: **2016**

Miejsce realizacji: **Warszawa, Polska**

Organ przyznający fundusze na realizację projektu: **European Molecular Biology Organization**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator**

9. Rodzaj projektu: **Networking Grant, EMBO (4 800 Euro)**

Tytuł projektu: **EMBO workshop on Computational Biology II**

Czas realizacji: **2016**

Miejsce realizacji: **Warszawa, Polska**

Organ przyznający fundusze na realizację projektu: **European Molecular Biology Organization**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator**

8. Rodzaj projektu: **Networking Grant, EMBO (15 000 Euro)**

Tytuł projektu: **YSF, EMBO**

Czas realizacji: **2015**

Miejsce realizacji: **Warszawa, Polska**

Organ przyznający fundusze na realizację projektu: **European Molecular Biology Organization**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator**

7. Rodzaj projektu: **Small grant, Visegrad (6 000 Euro)**

Tytuł projektu: **3rd Summer School in Molecular Biophysics and Systems Biology**

Czas realizacji: **2015**

Miejsce realizacji: **Nover Hrad, Czechy**

Organ przyznający fundusze na realizację projektu: **Visegrad fund**

Charakter udziału habilitanta w projekcie: **partner**

6. Rodzaj projektu: **Networking Grant, EMBO (3 000 Euro)**

Tytuł projektu: **EMBO workshop on Computational Biology I**

Czas realizacji: **2015**

Miejsce realizacji: **Biebrza, Polska**

Organ przyznający fundusze na realizację projektu: **European Molecular Biology Organization**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator**

5. Rodzaj projektu: **Small grant, Visegrad (6 000 Euro)**

Tytuł projektu: **2rd Summer School in Molecular Biophysics and Systems Biology**

Czas realizacji: **2014**

Miejsce realizacji: **Nover Hrady, Czechy**

Organ przyznający fundusze na realizację projektu: **Visegrad fund**

Charakter udziału habilitanta w projekcie: **partner**

4. Rodzaj projektu: **Networking Grant, EMBO** (3 000 Euro)

Tytuł projektu: **Biophysical Society Meeting**

Czas realizacji: **2014**

Miejsce realizacji: **Wydział Chemii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **European Molecular Biology Organization**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator**

3. Rodzaj projektu: **DUN 782/P-DUN/2014** – (47 000 PLN)

Tytuł projektu: **Biophysical Society Meeting**

Czas realizacji: **2014**

Miejsce realizacji: **Wydział Chemii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **MNiSW**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator ,we współpracy z Polskim Towarzystwem Biofizycznym**

2. Rodzaj projektu: **Biophysical Society grant** (10 000 USA)

Tytuł projektu: **Biophysical Society Meeting, Warsaw, Poland – main organizer**

Czas realizacji: **2013**

Miejsce realizacji: **Wydział Chemii, Uniwersytet Warszawski**

Organ przyznający fundusze na realizację projektu: **Biophysical Society, USA**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator**

1. Rodzaj projektu: **Grant konferencyjny**

Tytuł projektu: **Entanglement in biology; how nature controls the topology of proteins and DNA**

Czas realizacji: **2012**

Miejsce realizacji: **Banff, Kanada**

Organ przyznający fundusze na realizację projektu: **Banff International Research Station (BIRS), Banff, Kanada**

Charakter udziału habilitanta w projekcie: **kierownik / główny organizator**

B) Nagrody

5. **Grant Ideas Plus** (2 615 024 PLN) – MNiSW. Grant na realizację projektu ERC Starting Grant, który otrzymał najwyższą ocenę A w klasyfikacji ERC, jednakże z powodu ograniczeń budżetowych nie został przez ERC sfinansowany.

4. Nazwa nagrody: **Outstanding Women in Science**

Rok przyznania: **2016**

Nazwa organu przyznającego nagrodę: **Fundacja Roberta Boscha z nominacji FNP**

3. Nazwa nagrody: **CBSB14 Outstanding Young Researcher Award**

Rok przyznania: **2014**

Nazwa organu przyznającego nagrodę: **Computational Biophysical on Systems Biology Group**

2. Nazwa nagrody: **EMBO Installation Grant**

Rok przyznania: **2013**

Nazwa organu przyznającego nagrodę: **European Molecular Biology Organization (EMBO)**

Określenie tytułu z jakiego została przyznana nagroda: **grant badawczy**

1. Nazwa nagrody: **Nagroda za najlepszą pracę doktorską w IFPAN**

Rok przyznania: **2008**

Nazwa organu przyznającego nagrodę: **Instytutu Fizyki, Polska Akademia Nauk**

Określenie tytułu z jakiego została przyznana nagroda: **Nagroda za najlepszą pracę doktorską**

C) Dane bibliometryczne

Sumaryczny Impact Factor: **183.86**

Całkowita liczba cytowań według bazy danych Web of Science (z uwagą, iż baza ta nie wyszukuje poprawnie wszystkich cytowań moich prac z powodu zmiany nazwiska i polskiej litery „ł” w nazwisku): **833**

Całkowita liczba cytowań według bazy danych google scholar wynosi: **1113**

Indeks Hirscha (Web of Science): **16**

Liczba wszystkich publikacji (Web of Science): **36**

Liczba publikacji jako „corresponding author”: **8**

Liczba publikacji jako pierwszy autor: **16**

D) Wykaz i omówienie pozostałych publikacji

Publikacja z zakresu bioinformatyki i fizyki statystycznej

1. J. I. Sulkowska*, F. Marcos*, T. Hwa, J.N. Onuchic,
 “*Genomics Aided Structure Prediction (GASP)*”,
 Proc. Natl. Acad. Sci. USA (2012), 109(26): 10340-5.

Publikacje z zakresu biofizyki

2. Li Sun, Jeffrey K. Noel, J. I. Sulkowska, Herbert Levine, José N. Onuchic,
 “*Connecting Thermal and Mechanical Protein (Un)folded Landscapes*”,
 Biophys J. (2014) 16; 107(12):2941-52.

3. A. Valbuena, J. Oroz, R. Hervás, A. M. Vera, D. Rodrigues, A. Menedez, J. I. Sulkowska, M. Cieplak and M. Carrion-Vazquez,
 “*On the remarkable mechanostability of scaffoldins and the mechanical clamp motif*”,
 Proc. Natl. Acad. Sci. (2009) 106, 13791.

4. M. Sikora*, J. I. Sulkowska* and M. Cieplak,
 “*Mechanical strength of 17 134 model proteins and cysteine slipknots motif*”,
 PLoS Comput. Biol. (2009) 5, e1000547.

5. M. Sikora, J.I. Sulkowska, B.S. Witkowski, M. Cieplak,
 “*BSDB: the biomolecule stretching database*”,
 Nucleic Acids Res. (2011) 39:D443-50.

6. M. Cieplak, J. I. Sulkowska,

"Tests of the Structure-Based Models of Proteins",
Act. Phys Polonica A (2009) 115, 441.

7. J. I. Sulowska, A. Kloczkowski, T. Z. Sen, M. Cieplak and R. L. Jernigan,
"Predicting the Order in Which Contacts Are Broken during Single Molecule Protein Stretching Experiments",
Proteins: Structure, Function, and Bioinformatics (2008) 71, 45-60.

8. J. I. Sulowska, M. Cieplak,
"Selection of optimal variants of Go-like models of proteins through studies of stretching",
Biophys. J. (2008) 95, 3174.

9. J. I. Sulowska, M. Cieplak,
"Mechanical Stretching of proteins - A theoretical survey of the Protein Data Bank",
J. Phys. Cond. Mat. (2007) 19, 283201.

10. J. I. Sulowska, P. Sulowski, P. Szymczak, M. Cieplak,
"Tightening of knots in proteins",
Phys. Rev. Lett. (2008) 100, 058106.

11. M. Cieplak, J. I. Sulowska
"Thermal unfolding of proteins",
J. Chem. Phys. (2005) 123, 194908.

12. J. I. Kwiecińska (moje nazwisko panięskie), M. Cieplak,
"Chirality and protein folding",
J. Phys. Cond. Mat. (2005) 17, S1565.

13. M. Atakhorrami, J. I. Sulowska, K. M Addas, G. Koenderink, J. X Tang, A. J. Levine, F. C. MacKintosh, C. F. Schmidt,
"Correlated fluctuation of microparticles in viscoelastic solutions: quantitative measurement of materials properties by microrheology in the presence of optical traps",
Phys. Rev. E, 73, 061501 (2006).

Monografie, publikacje konferencyjne i inne

14. M. Cieplak, J. I. Sulowska
"Structure based models of biomolecules: stretching of proteins, dynamics of knots, hydrodynamic effects, and indentation of virus capsids",
Springer, A. Koliński, Chapter 8, New York, pp. 179-208, DOI 10.1007/978-1-4419-6889-0 (2011).
• rozdział w książce

15. J. I. Sulowska, P. Sulowski, P. Szymczak, M. Cieplak,
"Stretching the knotted protein YibK and its unknotted constructs",
Proceedings of the conference on "Knots and soft-matter physics", Kyoto University, Japan, (2009).

16. M. Cieplak, Sz. Niewieczerzal, J. I. Sulowska, P. Szymczak,
"Stretching of biomolecules in structure based models",
Longmans Orient, Bangalore, India (2009).

17. M. Cieplak, J. I. Sulowska,
„Rozciąganie molekul bialek – porównanie ich własności mechanicznych",
Kosmos, 55, 4 (2006).

Opis osiągnięć przedstawionych w powyższych publikacjach

Moje pozostałe osiągnięcia przedstawione są w publikacjach [1-13], jednym rozdziale w książce [14] oraz opracowaniach pokonferencyjnych i przeglądowych wymienionych powyżej. Prace te są poświęcone czterem zagadnieniom: i) przewidywaniu struktury białek [1]; ii) konstrukcji modeli gruboziarnistych oraz ich zastosowaniu do analizy własności mechanicznych białek [2-10]; iii) zastosowaniu modeli gruboziarnistych do analizy procesów termodynamicznych w białkach [11-12]; iv) własnościom wiskoelastycznym polimerów [13]. Otrzymane przeze mnie wyniki dotyczące tych zagadnień podsumowane są poniżej.

i) Przewidywanie struktury białek

Aby zrealizować swoją biologiczną funkcję białka muszą przyjąć ściśle określoną, trójwymiarową strukturę. Struktura ta może być wyznaczona doświadczalnie, jednakże wciąż znacznie więcej poznanych jest jedynie sekwencji aminokwasów tworzących białka, niż trójwymiarowych funkcjonalnych konfiguracji. Z powodu ogromnych kosztów związanych z doświadczalną identyfikacją struktur białek, zasadnym jest modelowanie takich konfiguracji teoretycznymi metodami bioinformatycznymi. Metody te całkiem dobrze przewidują struktury przestrzenne białek na podstawie znajomości ich homologicznej sekwencji oraz foldu, o czym świadczą m.in. wyniki międzynarodowych zawodów The Critical Assessment of protein Structure Prediction (CASP).

W pracy [1] skonstruowałam nową, efektywną hybrydową metodę DCA-fold do przewidywania struktury przestrzennej białek. Podstawą tej metody jest opracowanie zaskakująco skutecznej techniki do przewidywania oddziaływania między parami aminokwasów, na podstawie koewolucji par aminokwasów w sekwencji białek homologicznych, analizowanej technikami DCA („Direct Coupling Analysis”). Mimo że podobne techniki bazujące na koewolucji aminokwasów były już wcześniej stosowane, dopiero metody DCA pozwoliły na przewidzenie od 50 do 300 nielokalnych kontaktów z dokładnością 70-80%. W pracy [1] stworzyłam metodę hybrydową DCA-fold, opartą o model gruboziarnisty oraz technikę DCA, do przewidywania nielokalnych oddziaływań. Ponadto wyprowadziłam potencjały statystyczne do opisu charakteru oddziaływań dalekozasięgowych przewidzianych techniką DCA. Strukturę lokalną przewidziałam na podstawie potencjałów statystycznych dla kątów (ψ , ϕ) oraz korzystając z graficznych metod obrazowania macierzy kontaktów. Na podstawie skonstruowanego modelu przewidziałam struktury białek w zakresie 1-3 Å (RMSD od stanu krystalicznego) dla białek jednodomenowych o średniej liczbie 150 aminokwasów.

Warto zauważyć, iż praca [1] jest jedną z dwóch pierwszych w tej dziedzinie, a metoda DCA jest obecnie szeroko stosowana i wciąż rozwijana. Praca [1] została przeczytana ponad 9000 razy według bazy danych PNAS, a zacytowana od roku 2013 ponad 70 razy.

ii) Dynamika molekularna, mechaniczne rozwijanie i własności mechaniczne białek

Drogi jakimi zwijają i rozwijają się białka, mimo że w dużym stopniu poznane na podstawie analizy procesów termodynamicznych, dla większości białek nadal nie są wyznaczone. Proces zwijania i rozwijania białek może odbywać się w skalach czasowych rzędu od milisekund do minut. Symulowanie spontanicznego zwijania białka w czasie minut niestety jest obecnie poza zasięgiem wydajności komputerów. Alternatywnym podejściem do poznania mechanizmów zwijania/rozwijania białek są techniki ich mechanicznego rozciągania, których modelowanie w symulacjach komputerowych pozwala na skrócenie czasów osiągnięcia danego stanu o kilka rzędów wielkości. Ponadto doświadczalnie procesy rozciągania mogą być przeprowadzone dla pojedynczej molekuly białka.

Otrzymane przeze mnie wyniki określające krajobraz energetyczny białek, uzyskane na podstawie symulacji ich mechanicznego rozwijania, przedstawione są w cyklu prac [2-10]. W ramach tego cyklu skonstruowałam optymalny model gruboziarnisty typu Go do badania mechanicznego rozwijania białek, scharakteryzowałam przy jego pomocy własności mechaniczne 17134 białek zdeponowanych w PDB, oraz wyłoniłam najbardziej „oporne” mechanicznie białka. Optymalny model Go wyłoniłam spośród 54 skonstruowanych przeze mnie modeli, którymi zbadałam własności mechaniczne kilkunastu białek, a następnie porównałam je z danymi doświadczalnymi. Wyniki tych badań są zawarte w publikacjach [6,8]. W publikacji [10] dokonałam pierwszego na świecie teoretycznego, przekrojowego przeglądu własności mechanicznych białek, obejmującego 17134 struktury. Na podstawie otrzymanych wyników wyłoniłam 134 najbardziej odporne mechanicznie białka oraz wyznaczyłam kształt ich „mechanicznych imadeł”, czyli motywów geometrycznych odpowiedzialnych za dużą stabilność struktury białka. W pracach [8,9] ulepszyłam opracowany model i rozszerzyłam przegląd o kolejne tysiąc białek. Ponadto w pracy [8] zidentyfikowałam najbardziej mechanicznie odporne białko spośród tych o znanej strukturze przestrzennej, a w pracy [3] wraz z grupą doświadczalników wykazałam poprawność otrzymanych wyników doświadczalnie. W pracy [8] zidentyfikowałam także nowy geometryczny motyw, który jest odpowiedzialny za wyjątkowo duży opór mechaniczny. Wyżej wymienione prace doprowadziły do stworzenia przeze mnie internetowej bazy danych charakteryzującej własności mechaniczne białek o znanej strukturze przestrzennej „Bio-molecule Stretching Database” (BSDB), która jest dostępna pod adresem <http://jowisz.ifpan.edu.pl/BSDB/>. Baza BSDB jest obecnie powszechnie używana przez teoretyków i doświadczalników do określenia własności mechanicznych białek.

W pracy [7] zbudowałam model hybrydowy do analizy wszystkich zmian konformacyjnych w dużych biomolekułach, których struktury w stanie przejściowym są słabo dostępne ze względu na długie czasy ich symulowania, i których analiza w modelach pełnoatomowych jest obecnie poza zasięgiem nawet najszybszych komputerów. Otrzymany model wykorzystuje stany przejściowe wyznaczone podczas mechanicznego rozwijania w modelu Go, a następnie przewiduje kolejne możliwe rozwinięcia białka na podstawie amplitud fluktuacji wyznaczonych modelem Gaussian Network Model (GNM).

Publikacja [2] poświęcona jest zastosowaniu techniki szczypiec optycznych do analizy mechanicznego rozwijania białek. Niesamowity postęp w ostatnich kilku latach w rozdzielczości techniki szczypiec optycznych umożliwia obecnie pomiar mechanicznego rozwijania niemalże na poziomie termicznych fluktuacji, które są odpowiedzialne za zrywanie pojedynczych wiązań wodorowych w białku. W pracy [2] wykazałam różnice pomiędzy interpretacją danych doświadczalnych i teoretycznych (otrzymanych w symulacjach komputerowych). W pracy tej opracowałam także teoretyczny model do analizy danych doświadczalnych otrzymanych w przełomowym doświadczeniu ilustrującym zastosowanie szczypiec optycznych do wyznaczania ścieżek zwijania białka, wykonanym w grupie prof. Riefa (Technische Universität Muenchen, Germany), autora słynnej pracy „Full distance-resolved folding energy landscape of one single protein molecule”, PNAS (2010), 107, 5, 2013-2018.

W pracy [10] na podstawie symulacji komputerowych dwudziestu białek o nietrywialnej topologii wykazałam, iż proces mechanicznego rozwijania białek zawężonych jest znacząco różny od zawężonych homopolimerów. Ponadto pokazałam, iż proces zaciskania węzła (podczas rozciągania białka) odbywa się skokowo, a rozkład czasów może być opisany skokami Levy’ego. Analizując geometrię białka wykazałam, iż miejsca skoków są skorelowane z ostrymi zakrętami w łańcuchu białkowym, a węzeł zaciska się w natywnym obszarze. Wyniki tych badań zostały obecnie potwierdzone przez prace teoretyczne stosujące modele gruboziarniste prowadzone przez prof. Szymczaka (Uniwersytet Warszawski), wyniki symulacji w modelach pełnoatomowych otrzymane przez prof. Dziubiella (Humboldt-University of Berlin, Germany), oraz wyniki doświadczalne otrzymane przez prof. Riefa (Technische Universität Muenchen, Germany).

iii) Zastosowanie modeli gruboziarnistych do analizy procesów termodynamicznych w białkach

W pracach [11,12] zbadalam wpływ zastosowania uproszczonego modelu bazujacego na stanie natywnym bialka na mechanizmy spontanicznego zwijania, rozwijania i mechanicznego rozwijania bialek oraz porownalam je do wynikow pochodzacych z symulacji pelnoatomowych lub badan doswiadczalnych. W pracy [11], na podstawie wszechstronnych symulacji procesow rozwijania bialek, okreslilam definicje czasow rozwijania oraz pokazalam, ze wraz z obnizaniem temperatury ich wartosc rośnie szybciej niz wynikałoby to z prawa Arrheniusa. Ponadto analizujac krajobraz formowania i pękania kontaktow pokazalam, iz ponizej charakterystycznej temperatury dlugie czasy zwijania pojedynczych elementow bialka antykorelują się z krótkimi czasami rozwijania tych samych struktur (podczas symulacji procesow rozwijania). W pracy [12] wykazalam, ze kryteria: wartosc sredniego kwadratowego odchylenia od stanu natywnego oraz liczba natywnych kontaktow są niewystarczajace aby poprawnie ocenic zwinięcie bialka w modelu gruboziarnistym zbudowanym wyłacznie z węgli C α . Kryteria te nie są czule np. na złą (odwrotną niz natywna) chiralność helisy. W tej samej pracy pokazalam, iz rozwiązaniem tego problemu jest wprowadzenie np. dodatkowego potencjalu chiralnościowego do hamiltonianu w modelu gruboziarnistym lub uwzględnienie w modelu gruboziarnistym węgli C β (pierwszy atom w grupie bocznej aminokwasu) wymuszajacych natywną skrętność bialka.

iv) Właściwości wiskoelastyczne polimerów

W pracy [13] zbadalam fluktuacje mikroczastek w wiskoelastycznym roztworze polimerow, wykonujac serie pomiarow doswiadczalnych za pomoca optycznych szczyptic. Ponadto w pracy tej skonstruowalam i zaprogramowalam model do analizy korelacji fluktuacji dwuch mikroczasteczek. Obecnie pomiary za pomoca dwuch optycznych szczypticzykow są powszechnie prowadzone w ramach badania procesow rozwijania RNA i bialek.

Joanna Sulkowska