

Międzyuczelniany Wydział Biotechnologii
Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego

Piotr Sławomir Karaś

Molekularne podstawy ewolucji małych białek szoku cieplnego w *Erwiniaceae*

Molecular basis of small heat shock protein
evolution in *Erwiniaceae*

Praca przedstawiona

Radzie Dyscypliny Biotechnologia Uniwersytetu Gdańskiego

celem uzyskania stopnia doktora

w dziedzinie nauk ścisłych i przyrodniczych

w dyscyplinie naukowej biotechnologia

Promotor: prof. dr hab. Krzysztof Liberek

Pracownia Biochemii Białek

Promotor pomocniczy: dr inż. Bartłomiej Tomiczek

Pracownia Biochemii Ewolucyjnej

Gdańsk 2023



NARODOWE CENTRUM NAUKI

Badania finansowane przez Narodowe Centrum Nauki w
ramach grantu OPUS 17 2019/33/B/NZ1/00352

*Dwubiałkowy system małych białek szoku termicznego jako
pierwsza linia obrony w odpowiedzi na stres u Enterobacterales*

Chciałbym serdecznie podziękować:

Mojemu promotorowi prof. Krzysztofowi Liberkowi

Za stworzenie możliwości rozwoju naukowego, przekazaną wiedzę, otwartość na dyskusje oraz okazane wsparcie i dobre rady

Mojemu promotorowi pomocniczemu dr. Bartłomiejowi Tomiczkowi

Za wprowadzenie w tajniki analiz bioinformatycznych i ewolucyjnych, cenną, konstruktywną krytykę i nieustający optymizm

Dr. Igorowi Obuchowskiemu

Za przekazane doświadczenie w pracy z białkami opiekuńczymi i cierpliwość w odpowiadaniu na pytania

Koleżankom i Kolegom z Pracowni Biochemii Białek, Pracowni Biologii Molekularnej i Pracowni Biochemii Ewolucyjnej

Za przyjazną, ciepłą atmosferę, gotowość do służenia radą i pomocą oraz niezliczone inspirujące dyskusje

Mojej Rodzinie i Przyjaciołom

Za ofiarowane wsparcie zarówno w dobrych jak i trudnych chwilach

Spis treści

1. Streszczenie	7
2. Abstract	9
3. Spis skrótów	10
4. Wstęp	11
4.1 Rola białek opiekuńczych w proteostazie	11
4.1.1 Białka Hsp60.....	13
4.1.2 Białka Hsp90.....	14
4.1.3 System Hsp70.....	14
4.1.4 Białka Hsp100.....	16
4.2 Małe białka szoku cieplnego	18
4.2.1 Struktura małych białek szoku cieplnego.....	19
4.2.2 Mechanizm działania i funkcja małych białek szoku cieplnego	22
4.3 Małe białka szoku cieplnego w <i>Enterobacterales</i>	24
4.3.1 Dwubiałkowy system sHsp w <i>Enterobacteriaceae</i>	26
4.3.2 Wtórnie pojedyncze lbpA w <i>Erwiniaceae</i>	27
5. Cel Pracy.....	28
6. Materiały i Metody.....	29
6.1 Materiały.....	29
6.1.1 Szczepy bakterii.....	29
6.1.2 Plazmidy	29
6.1.3 Startery oligonukleotydowe.....	29
6.1.4 Syntetyczne geny	30
6.1.5 Sekwencje aminokwasowe wykorzystane do analizy filogenetycznej i rekonstrukcji ancestralnej:.....	31
6.1.6 Sekwencje nukleotydowe wykorzystane do analizy selekcji naturalnej:.....	32
6.1.7 Białka	32
6.1.8 Pożywki	33
6.1.9 Antybiotyki	33
6.1.10 Bufory.....	33
6.2 Metody.....	34
6.2.1 Analiza filogenetyczna i rekonstrukcja ancestralna	34
6.2.2 Synteza genów <i>de novo</i>	35
6.2.3 Mutageneza miejscowo - specyficzna.....	35
6.2.4 Izolacja plazmidowego DNA.....	36

6.2.5 Przygotowanie komórek kompetentnych.....	36
6.2.6 Transformacja komórek kompetentnych plazmidowym DNA.....	36
6.2.7 Nadprodukcja i oczyszczanie białek	37
6.2.7.1 IbpAE.coli, domena ACD białka IbpA _{E. coli} , DnaK, DnaJ, GrpE, ClpB, lucyferaza ze znacznikiem polihistydynowym.....	37
6.2.7.2 Białka IbpA	37
6.2.7.3 Białko IbpB z <i>E. coli</i>	38
6.2.7.4 Domeny ACD białek AncA ₀ i AncA ₀ Q66H G109D	39
6.2.7.5 His ₆ – SUMO – peptyd C.....	41
6.2.8 Przygotowanie lizatu <i>E. coli</i>	41
6.2.9 Oznaczanie stężenia białek	42
6.2.10 Reaktywacja lucyferazy	42
6.2.11 Dynamiczne Rozpraszanie Światła (DLS).....	43
6.2.11.1 Analiza odwracalnej deoligomeryzacji sHsp w warunkach szoku cieplnego.....	43
6.2.11.2 Analiza zdolności sHsp do sekwestracji agregujących substratów	43
6.2.11.3 Analiza wpływu substytucji na tworzenie oligomerów sHsp	44
6.2.12 Interferometria Biowarstwowa (BLI)	44
6.2.12.1 Analiza oddziaływania sHsp i domen ACD z agregatami białkowymi	44
6.2.12.2 Analiza oddziaływania domen ACD z regionem CTE sHsp	45
6.2.12.3 Analiza inhibicji wiązania systemu Hsp70 do agregatów białkowych przez sHsp.....	45
6.2.13 Analiza selekcji naturalnej.....	46
6.2.14 Analiza strukturalna i dynamika molekularna	46
7. Wyniki.....	48
7.1 Porównanie zdolności współczesnych małych białek szoku cieplnego z <i>Erwiniaceae</i> i <i>Enterobacteriaceae</i> do wspomagania dezagregacji	48
7.2 Analiza filogenezy białek IbpA w rzędzie <i>Enterobacterales</i>	49
7.3 Rekonstrukcja ancestralnej sekwencji aminokwasowej białek AncA ₀ i AncA ₁	52
7.4 Charakterystyka biochemiczna rekonstruowanych białek AncA ₀ i AncA ₁	53
7.4.1 Analiza zdolności białek AncA ₀ i AncA ₁ do odwracalnej deoligomeryzacji w warunkach szoku cieplnego	53
7.4.2 Analiza zdolności białek AncA ₀ i AncA ₁ do wiązania agregatów białkowych	54
7.4.3 Analiza zdolności białek AncA ₀ i AncA ₁ do sekwestracji agregującego substratu .	57
7.4.4 Analiza zdolności białek AncA ₀ i AncA ₁ do wspomagania dezagregacji i ponownego zwijania białek.....	58
7.5 Analiza zależności właściwości rekonstruowanych białek od niepewności statystycznej wynikającej z rekonstrukcji ancestralnej	60

7.6 Analiza selekcji naturalnej pomiędzy węzłami AncA ₀ i AncA ₁	64
7.7 Identyfikacja kluczowych mutacji odpowiadających za różnice w aktywności białek AncA ₀ i AncA ₁	66
7.8 Analiza znaczenia reszt 66 i 109 dla różnic w aktywności białek AncA ₀ i AncA ₁ na poziomie molekularnym	72
7.8.1 Lokalizacja reszt 66 i 109 w strukturze białka sHsp	72
7.8.2 Analiza wpływu substytucji na pozycjach 66 i 109 na siłę oddziaływania między domeną ACD, a peptydem C <i>in vitro</i>	73
7.8.3 Analiza wpływu substytucji na pozycjach 66 i 109 na stan oligomeryczny sHsp..	76
7.8.4 Analiza wpływu substytucji na pozycjach 66 i 109 na wiązanie zagregowanego substratu przez domenę ACD <i>in vitro</i>	77
7.9 Analiza znaczenia pozycji 66 i 109 we współczesnych białkach IbpA z <i>E. coli</i> i <i>E. amylovora</i>	80
8. Dyskusja.....	87
9. Bibliografia	93

1. Streszczenie

Małe białka szoku cieplnego (sHsp) są w stanie wiązać agregujące substraty, sekwestrując je w konformacji bliskiej do natywnej w złożonych kompleksach sHsp – substrat, zapobiegając tym samym niekontrolowanej dalszej agregacji. Bliska natywnej konformacja sekwestrowanych substratów jak i niewielkie rozmiary tworzonych kompleksów ułatwiają późniejszy proces dezagregacji i przywracania konformacji natywnej, prowadzony przez system Hsp70 i dezagregazy Hsp100. Inicjacja tego procesu wymaga jednak wyparcia przez system Hsp70 białek sHsp związanych do substratu, co wymusza kompromis między wydajnym wiązaniem i sekwestracją substratu przez białka sHsp, a ich zdolnością do stymulacji dezagregacji prowadzonej przez system Hsp70 i Hsp100.

Większość bakterii należących do rzędu *Enterobacterales* (w tym rodzina *Enterobacteriaceae*) posiada system dwóch współpracujących białek sHsp – lbpA i lbpB, powstałych na skutek duplikacji genu kodującego ancestralne, pierwotnie pojedyncze lbpA. lbpA wchodzące w skład systemu dwubiałkowego silnie oddziałuje z substratem i posiada silną aktywność sekwestrazy, podczas gdy lbpB, samodzielnie dużo słabiej wiążące substraty, ułatwia dysocjację sHsp od substratu podczas procesu dezagregacji, umożliwiając efektywną stymulację tego procesu przez sHsp nawet przy niższych stężeniach Hsp70. W rodzinie *Erwiniaceae* lbpB zostało utracone, a wtórnie pojedyncze lbpA wykształciło zdolność do efektywnej stymulacji dezagregacji prowadzonej przez system Hsp70 i Hsp100 na poziomie zbliżonym do systemu dwubiałkowego z rodziny *Enterobacteriaceae* mimo braku partnera.

W ramach tej pracy podjąłem próbę wyjaśnienia, w jaki sposób wtórnie pojedyncze lbpA z *Erwiniaceae* nabyło swoją nową funkcjonalność. W tym celu przeprowadziłem rekonstrukcję historii ewolucyjnej białka lbpA w *Enterobacterales*, pokazując, że charakterystyczna aktywność wtórnie pojedynczego lbpA z *Erwiniaceae* wykształciła się już w ostatnim wspólnym przodku tej rodziny, równoległe do utraty paralogicznego białka lbpB, najprawdopodobniej pod wpływem działania selekcji pozytywnej. Porównując sekwencje i właściwości biochemiczne rekonstruowanych białek ancestralnych, zidentyfikowałem dwie substytucje aminokwasowe (Q66H i G109D), które umożliwiły temu białku nabycie zdolności do efektywnej stymulacji dezagregacji i przywracania konformacji natywnej sekwestrowanych polipeptydów. Pokazałem następnie, że substytucje te spowodowały osłabienie oddziaływania między domeną α -krystaliny, a regionem końca karboksylowego białka lbpA, jak również oddziaływania między tą domeną, a substratem. Osłabienie tych oddziaływań ułatwiło wypieranie białek lbpA z kompleksu z substratem przez system Hsp70, zwiększając wydajność stymulacji dezagregacji sekwestrowanego substratu. Pokazałem również, że reszty aminokwasowe znajdujące się na zidentyfikowanych

kluczowych pozycjach odgrywają istotną rolę w różnicach funkcjonalnych między współczesnymi IbpA z rodzin *Erwiniaceae* i *Enterobacteriaceae*.

2. Abstract

Small heat shock proteins (sHsp) bind aggregating protein substrates and sequester them in nearly – native conformation in sHsp - substrate assemblies, preventing uncontrolled protein aggregation. Nearly – native conformation of the sequestered substrates combined with relatively small size of sHsp – substrate assemblies facilitates later disaggregation and refolding of those substrates by Hsp70 system and Hsp100 disaggregases. Initiation of the disaggregation process requires Hsp70 system replacing sHsp bound to the substrate, which leads to the tradeoff between sHsp ability to bind and sequester substrates (sequestrase activity) and their ability to stimulate Hsp70 – Hsp100 - mediated disaggregation and refolding.

Majority of *Enterobacterales* (including family *Enterobacteriaceae*) contains a system of two cooperating sHsps – lbpA and lbpB, which most likely evolved by a duplication of ancestral, primarily single lbpA. lbpA protein from that two – protein system strongly interacts with the substrate and exhibits strong sequestrase activity. lbpB, on the other hand, on its own interacts weakly with most substrates, but can facilitate sHsp displacement from sHsp – substrate assemblies during disaggregation, enabling the sHsps to efficiently stimulate this process even at lower Hsp70 concentrations. In *Erwiniaceae* family the lbpB protein was lost and the secondarily single lbpA evolved the ability to facilitate Hsp70 – Hsp100 – mediated disaggregation and refolding with effectiveness similar to two – protein sHsp system, even in the absence of the lbpB partner.

In this study I investigated how the secondarily single lbpA from *Erwiniaceae* developed its new functionality. I reconstructed the evolutionary history of lbpA in *Enterobacterales*, showing that the characteristic activity of *Erwiniaceae* secondarily single lbpA developed in the last common ancestor of that family, in parallel to the loss of paralogous protein lbpB, most likely due to positive selection. By comparing the sequences and biochemical properties of reconstructed ancestral proteins, I identified two substitutions (Q66H and G109D) that enabled *Erwiniaceae* lbpA to much more effectively stimulate Hsp70 – Hsp100 - mediated disaggregation and refolding of sequestered proteins. I have shown that those substitutions weakened interactions between ACD domain and CTE region of lbpA, as well as between lbpA and substrate, which in turn facilitated lbpA displacement from the substrate by Hsp70, increasing the effectiveness of the disaggregation stimulation. Moreover, amino acid residues on positions homologous to the two identified play a key role in functional differences between extant sHsp from families *Erwiniaceae* and *Enterobacteriaceae*.

3. Spis skrótów

- **AAA+** - (ang. **A**TP-ases **a**ssociated with various cellular **a**ctivities) – ATP – azy związane z różnymi aktywnościami komórkowymi
- **Domena ACD** – (ang. **a**lpha – **c**rystallin **d**omain) domena α-kryształinowa
- **ADP** – (ang. **a**denosine **d**iphosphate) difosforan adenozyiny
- **ATP** – (ang. **a**denosine **t**riphosphate) trifosforan adenozyiny
- **BLI** – (ang. **b**io**l**ayer **i**nterferometry) interferometria biowarstwowa
- **BSA** – (ang. **b**ovine **s**erum **a**lbumin) albumina osocza wołowego
- **Region CTE** – (ang. **C**-**t**erminal **e**xtention) region na końcu karboksylowym białka
- **DLS** – (ang. **D**ynamic **L**ight **S**cattering) dynamiczne rozpraszanie światła
- **DTT** – 1,4 – **d**itiotreitol
- **EDTA** – (ang. **e**thylene**d**iamine**t**etra**a**cetic **a**cid) kwas wersenowy (etylenodiaminotetraoctowy)
- **His₆** – znacznik heksahistydynowy
- **Hsp** – (ang. **h**eat **s**hock **p**roteins) białka szoku cieplnego
- **IPTG** - **i**zopropyl-β-D-**t**iogalaktopiranozyd
- **Domena NBD** – (ang. **n**ucleotide **b**inding **d**omain) domena wiążąca nukleotyd
- **Ni - NTA** – (ang. **n**itrylo**t**riacetic **a**cid) kwas nitrylowotrioctowy ze związanymi jonami Ni²⁺
- **Region NTE** – (ang. **N**-**t**erminal **e**xtention) region na końcu aminowym białka
- **OD₆₀₀** – (ang. **o**ptical **d**ensity) rozpraszanie światła lub absorbanca mierzona przy długości fali 600 nm
- **Domena SBD** – (ang. **s**ubstrate **b**inding **d**omain) domena wiążąca substrat
- **SDS** – (ang. **s**odium **d**odecyl **s**ulphate) Dodecylosiarczan sodu
- **SDS-PAGE** – (ang. **s**odium **d**odecyl **s**ulphate – **p**oliacrylamide **g**el **e**lectrophoresis) elektroforeza poliakrylamidowa w warunkach denaturujących
- **sHsp** – (ang. **s**mall **H**eat **S**hock **P**roteins) małe białka szoku cieplnego
- **Tris** – **T**ris(hydroksymetylo)aminometan

4. Wstęp

4.1 Rola białek opiekuńczych w proteostazie

Białka odgrywają wiele kluczowych ról w organizmach żywych, m. in. jako enzymy, białka strukturalne, regulatorowe czy receptory. Aby jednak mogły one pełnić swoją fizjologiczną funkcję, muszą znajdować się w określonej konformacji, zwanej konformacją natywną. Zazwyczaj odpowiada ona globalnemu minimum energetycznemu przestrzeni konformacyjnej pojedynczej cząsteczki danego białka (Ryc. 1) (Jahn i Radford 2008; Vabulas i in. 2010).

Nowo syntetyzowane białka zwijają się do konformacji natywnej, przyjmując szereg konformacji o coraz mniejszej energii swobodnej. Podczas tego procesu spadek entropii towarzyszący zwijaniu białka równoważony jest wzrostem entalpii na skutek oddziaływań stabilizujących strukturę natywną (Dobson 2004; Vabulas i in. 2010).

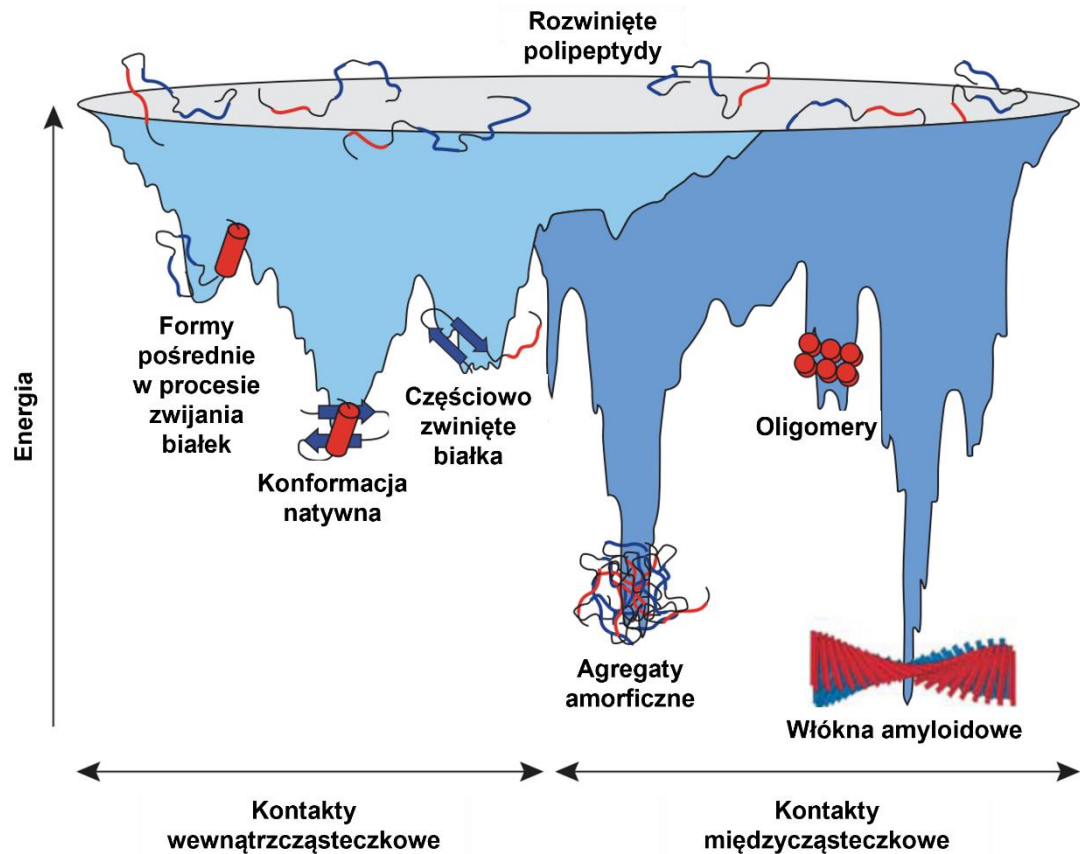
Zwijanie niektórych białek, zwłaszcza o większej masie czy szczególnie skomplikowanej strukturze, zachodzi poprzez szereg metastabilnych, nienatywnych konformacji pośrednich, stanowiących lokalne minima energetyczne przestrzeni konformacyjnej (Ryc. 1) (Jahn i Radford 2008; Vabulas i in. 2010).

Innym źródłem białek o nienatywnej konformacji jest utrata konformacji natywnej przez poprawnie zwinięte białka. Oddziaływania stabilizujące konformację natywną mogą być zaburzone przez czynniki stresowe, takie jak wysoka temperatura czy zmiany pH. Podobnie modyfikacje chemiczne aminokwasów, jak deaminacja czy utlenianie, np. na skutek stresu oksydacyjnego, mogą również powodować destabilizację konformacji natywnej (Moreau i King 2012). Prowadzi to do całkowitej lub częściowej denaturacji białek, które mogą następnie zwijać się konformacji nienatywnych (Ryc. 1).

Na denaturację białek na skutek szoku cieplnego narażone są w szczególności organizmy jednokomórkowe czy osiadłe (jak np. rośliny). W przypadku zwierząt, które dzięki wysokiej mobilności i złożonym mechanizmom termoregulacyjnym są w mniejszym stopniu narażone na szok cieplny, relatywnie większą rolę odgrywa denaturacja białek na skutek chronicznego stresu (np. oksydacyjnego) czy mutacji obniżających stabilność białek.

Białka znajdujące się w konformacji nienatywnej, czy to jako metastabilny etap pośredni fizjologicznego procesu zwijania białek, czy też na skutek utraty konformacji natywnej na skutek stresu, często zawierają eksponowane regiony o charakterze hydrofobowym, które w konformacji natywnej znajdowałyby się w głębi struktury trzyczłonowej. Takie eksponowane regiony znajdujące się na różnych cząsteczkach nieprawidłowo zwiniętych białek mogą oddziaływać ze sobą, prowadząc

do procesu agregacji – łączenia się białek w stabilne, nierozpuszczalne struktury, zwane agregatami (Vabulas i in. 2010). Ryzyko zajścia tego procesu jest dodatkowo potęgowane przez wysokie stężenie białek w komórkach, prowadzące do częstych kontaktów międzycząsteczkowych (efekt tzw. tłoku molekularnego, ang. *molecular crowding*) (Ellis i Minton 2006; Vabulas i in. 2010). Agregacja białek może prowadzić do powstawania amorficznych agregatów lub wysoce ustrukturyzowanych włókien amyloidowych, stabilizowanych przez oddziaływania między uszeregowanymi strukturami β (Ryc. 1) (Jahn i Radford 2008; Vabulas i in. 2010; Dobson 2004).



Ryc. 1 Schemat krajobrazu energetycznego przestrzeni konformacyjnej białka: *Jasnoniebieski obszar reprezentuje przestrzeń konformacyjną pojedynczej cząsteczki białka, wynikającą z kontaktów wewnątrzcząsteczkowych. Ciemnoniebieski obszar reprezentuje dodatkową przestrzeń konformacyjną, wynikającą z oddziaływań międzycząsteczkowych, potencjalnie promujących proces agregacji. Zaadaptowano z (Vabulas i in. 2010)*

W obu przypadkach niekontrolowany proces agregacji jest bardzo szkodliwy dla organizmu – białka uwięzione w agregatach nie są w stanie przyjąć konformacji natywnej i tym samym nie mogą pełnić swojej fizjologicznej funkcji. Co więcej, na powierzchni agregatów mogą także znajdować się eksponowane regiony hydrofobowe, które mogą służyć za centra nukleacji dla dalszego procesu agregacji. W przypadku tworzenia włókien amyloidowych, formy pośrednie prowadzące do ich powstawania często same mają charakter toksyczny dla błon komórkowych.

Agregacja białek związana jest z genezą wielu chorób, takich jak zaćma, w której agregacja białek soczewki oka prowadzi do jej zmętnienia czy choroby neurodegeneracyjne, np. Alzheimera i Parkinsona, którym towarzyszy odkładanie się złogów amyloidowych w ośrodkowym układzie nerwowym.

W związku z powyższym utrzymywanie stabilności proteomu, często określanej mianem homeostazy białkowej lub proteostazy, jest niezbędne dla prawidłowego funkcjonowania organizmów żywych. Kluczową rolę odgrywa w tym procesie grupa białek zwanych białkami opiekuńczymi (ang. *molecular chaperones*).

Definiuje się je jako białka, które oddziałują z nienatywnymi białkami, pomagając im osiągnąć ich konformację natywną lub stabilizują tę konformację, nie wchodząc przy tym w skład ostatecznej, funkcjonalnej struktury (Ellis 1997).

Białka opiekuńcze pomagają zachować proteostazę poprzez zapobiegające niekontrolowanej agregacji wiązanie nieprawidłowo zwiniętych białek, a także ułatwianie zwijania się nieprawidłowo lub częściowo zwiniętych białek do konformacji natywnej. Uczestniczą one również w procesie dezagregacji, uwalniając białka z agregatów, co umożliwia przywrócenie im konformacji natywnej (często również z pomocą białek opiekuńczych) lub ich degradację (Vabulas i in. 2010; Dobson 2004; Liberek, Lewandowska, i Ziętkiewicz 2008).

Niektóre białka opiekuńcze mogą pełnić również funkcje, które nie są bezpośrednio związane z utrzymywaniem proteostazy, jak rozwijanie i zwijanie białek podczas ich transportu przez kanały transbłonowe czy przenoszenie centrów żelazowo – siarkowych do enzymów (Lim i in. 2001; Ernst i in. 2017; Dutkiewicz i Nowak 2018).

Ze względu na fakt, że ekspresja wielu białek opiekuńczych ulega zwiększeniu w warunkach szoku cieplnego, wiele z nich określa się czasem jako białka szoku cieplnego (ang. *heat shock proteins*, *Hsp*).

Do białek opiekuńczych zalicza się grupa niespokrewnionych rodzin białek, z których najważniejsze to białka Hsp60, białka Hsp90, białka systemu Hsp70, białka Hsp100 i małe białka szoku cieplnego.

4.1.1 Białka Hsp60

Białka Hsp60, reprezentowane m.in. przez bakteryjne GroEL, funkcjonują jako oligomery w formie pierścienia złożonego z 7 monomerów z tzw. przykrywką utworzoną przez białka pomocnicze Hsp10 (np. bakteryjne GroES) (Tang i in. 2006; Vabulas i in. 2010).

Występują one w komórkach prokariotycznych, a także w mitochondriach i chloroplastach. Analogiczne funkcje w cytoplazmie komórek eukariotycznych pełnią białka TRiC/CCT. Ich struktura jest zbliżona do białek Hsp60, jednak tworzony przez

nie pierścien oligomeryczny składa się z 9 monomerów, a funkcję białka pomocniczego Hsp10 pełni helikalna struktura będąca integralną częścią właściwego białka (Meyer i in. 2003; Vabulas i in. 2010).

Białka Hsp60 wiążą niecałkowicie zwinięte białka i na skutek zmian konformacyjnych spowodowanych hydrolizą ATP zamykają je wewnątrz pierścienia oligomeru. Izoluje to zwijające się białka, zapobiegając niepożądanym oddziaływaniom międzycząsteczkowym, mogącym prowadzić do agregacji. Co więcej, ograniczona przestrzeń wnętrza oligomeru w połączeniu z wysoce hydrofilowym charakterem wnętrza pierścienia obniża bariery energetyczne między metastabilnymi konformacjami pośrednimi, a konformacją natywną, stymulując poprawne zwijanie się białek. Jest to tzw. efekt „klatki Affinsena” (ang. *Affinsen cage*) (Tang i in. 2006; Vabulas i in. 2010).

4.1.2 Białka Hsp90

Białka Hsp90, reprezentowane m. in. przez bakteryjne białko HtpG, funkcjonują jako dimery, mogące występować w dwóch konformacjach – otwartej i zamkniętej. Przejście między tymi konformacjami związane jest z cyklem hydrolizy ATP.

W odróżnieniu od pozostałych białek opiekuńczych, białka Hsp90 wiążą substraty znajdujące się na późnych etapach procesu zwijania białek, posiadające konformację zbliżoną do natywnej. Ich rola nie jest całkowicie poznana, ale polega głównie na modyfikacji i stabilizacji konformacji wiązanych substratów, m. in. regulując uporządkowane składanie i rozpad złożonych kompleksów białkowych. Odgrywają one kluczową rolę w końcowych etapach zwijania kinaz treoninowych i receptorów hormonów sterydowych, tworzeniu elementów szkieletu komórkowego czy kinetochorów, a w warunkach stresowych stabilizują termolabilne konformacje wielu białek, w tym regulatorów cyklu komórkowego.

Oddziaływanie białek Hsp90 z szeroką gamą substratów regulowane jest przez złożony system białek pomocniczych. (McClellan i in. 2007; Krukenberg i in. 2008; Pearl 2016).

4.1.3 System Hsp70

Białka Hsp70, w bakteriach reprezentowane przez DnaK, składają się z domeny wiążącej substrat - SBD (ang. *substrate binding domain*) i domeny wiążącej nukleotyd - NBD (ang. *nucleotide binding domain*). Domena SBD wykazuje powinowactwo do charakterystycznych hydrofobowych peptydów, eksponowanych przez niecałkowicie lub nieprawidłowo zwinięte białka (Mayer i Bukau 2005).

Wiązanie i hydroliza ATP przez domenę NBD powodują przechodzenie domeny SBD między konformacją otwartą, w której substrat może być szybko

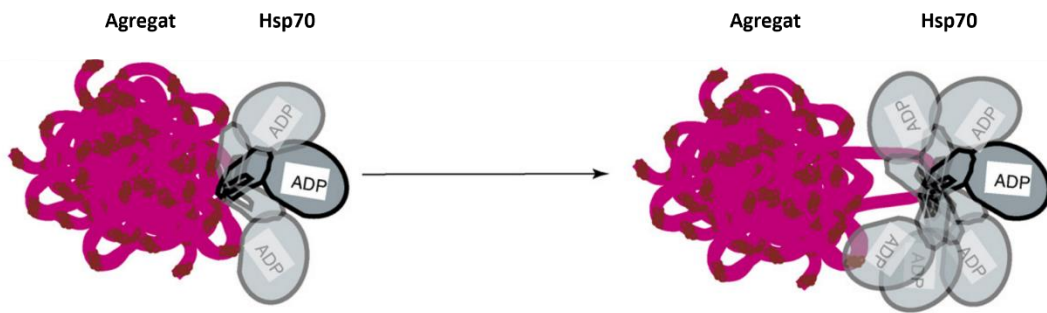
wiązany i łatwo dysocjować i konformacją zamkniętą, w której substrat jest uwięziony w kieszeni wiążącej. W odróżnieniu od białek Hsp60, zamknięciu w obrębie domeny SBD ulega jedynie niewielki fragment białka, a nie cała cząsteczka (Bertelsena i in. 2009; Mayer i Bukau 2005).

Poprawne funkcjonowanie białek Hsp70 wymaga współpracy z białkami pomocniczymi, razem z białkiem Hsp70 składającymi się na tzw. system Hsp70. Białka Hsp40 (nazywane również białkami J), reprezentowane m. in. przez bakteryjne białko DnaJ, rekrutują substraty do białka Hsp70 i stymulują jego aktywność ATP-azową. Są one odpowiedzialne za specyficzność substratową systemu (Mayer i Bukau 2005). Czynniki wymiany nukleotydów (białka NEF, z ang. *nucleotide exchange factor*) stymulują z kolei wymianę ADP związanego przez domenę NBD na ATP, umożliwiając wznowienie cyklu hydrolizy ATP. Do białek NEF należy kilka niespokrewnionych rodzin białek. W bakteriach są one reprezentowane przez białko GrpE (Mayer i Bukau 2005).

System Hsp70 bierze udział w początkowym etapie zwijania białek, przed przekazaniem substratów do białek Hsp60 lub Hsp90. Cykliczne wiązanie i dysocjacja regionów hydrofobowych przez system Hsp70 pozwala na ochronę zwijających się białek przed niekontrolowaną agregacją, umożliwiając ich poprawne zwijanie do konformacji natywnej.

W odróżnieniu od wyżej opisanych białek opiekuńczych, system hsp70, poza swoją rolą w procesie zwijania nowopowstałych białek, odgrywa również kluczową rolę w procesie dezagregacji – uwalniania białek z agregatów w celu umożliwienia ich powrotu do konformacji natywnej lub degradacji.

Proces ten najprawdopodobniej przebiega według mechanizmu określanego w literaturze anglojęzycznej jako „*entropic pulling*”. Bazuje on na fakcie, że białko Hsp70 związane do powierzchni agregatu posiada ograniczoną swobodę ruchu ze względu na zawadę steryczną. Dążenie do zwiększenia tej swobody ruchu, a co za tym idzie zwiększenia entropii układu, jest źródłem siły oddalającej białko od powierzchni agregatu. Siła ta, wywierana przez wiele systemów Hsp70, jest wystarczająca do zaburzenia struktury agregatu i uwolnienia z niego substratów wiązanych przez Hsp70 (Ryc. 2) (Goloubinoff i Rios 2007).



Ryc. 2 Mechanizm uwalniania substratu z agregatów przez Hsp70: *Hsp70* związane do powierzchni agregatu posiada ograniczoną swobodę ruchu, ze względu na zawadę steryczną wywoływaną bliskością agregatu i potencjalnie sąsiednich związanych białek *Hsp70*. Dążenie do zwiększenia entropii przez zwiększenie swobody ruchu jest źródłem siły oddalającej *Hsp70* od agregatu i tym samym rozluźniającej strukturę agregatu. Półprzezroczyste kształty reprezentują potencjalne pozycje, które mogą być zajęte przez białko *Hsp70* związane do agregatu. Zaadaptowano z (Goloubinoff i Rios 2007).

W procesie dezagregacji i przywracania uwolnionych substratów do konformacji natywnej system *Hsp70* współpracuje z białkami *Hsp100*. Współpraca ta jest szczególnie istotna w przypadku białek o konformacji dalekiej od natywnej, zwłaszcza wysoce stabilnych, nienatywnych struktur. (Liberek, Lewandowska, i Ziętkiewicz 2008).

4.1.4 Białka *Hsp100*

Białka *Hsp100*, reprezentowane m. in. przez bakteryjne ClpB i drożdżowe *Hsp104*, współpracują z systemem *Hsp70* w procesie dezagregacji. Należą one do klasy tzw. białek AAA+, czyli ATP-az związanych z różnymi aktywnościami komórkowymi (ang. *ATP - ases Associated with various cellular Activities*). Monomer białek *Hsp100* składa się z czterech domen – domeny N, zlokalizowanej na końcu aminowym, dwóch domen AAA, posiadających aktywność ATP-azy oraz regulatorowej domeny M (ang. *middle domain*), znajdującej się w środku pierwszej domeny AAA (Mogk, Kummer, i Bukau 2015).

Funkcjonują one jako heksamery w formie pierścienia. W jego centralnym porze znajdują się konserwowane pętle, zawierające aminokwasy o charakterze aromatycznym, będące w stanie wiązać substraty. Cykliczna hydroliza ATP przez domeny AAA poszczególnych monomerów prowadzi do zmian konformacyjnych w tych pętlach, powodujących przewlekanie związanego przez nie polipeptydu przez centralny por heksameru *Hsp100*. Przewlekanie to może zachodzić zarówno od końca karboksylowego, jak i aminowego, możliwe jest też przewleczenie struktury w formie pętli (Mogk, Kummer, i Bukau 2015).

Siła wywierana na przewlekany substrat może sięgać 20 pN i w połączeniu z dodatkowymi siłami ścinającymi generowanymi przez domenę N jest wystarczająca

do całkowitego rozwinięcia większości substratów. Umożliwia to nieprawidłowo zwiniętym białkom ponowne zwinięcie do konformacji natywnej (samodzielnie lub z pomocą innych białek opiekuńczych, np. systemu Hsp70) (Mogk, Kummer, i Bukau 2015).

Specyficzność substratowa i aktywność białek Hsp100 wymaga ścisłej regulacji przez system Hsp70. Konstytutywnie aktywne mutanty, pozbawione tej regulacji, są toksyczne dla komórek – najprawdopodobniej denaturują one również poprawnie zwinięte, natywne białka (Lipińska i in. 2013; Chamera i in. 2019).

Kiedy białko Hsp100 nie oddziałuje z systemem Hsp70, domena M hamuje aktywność ATP-azową domen AAA. Związane z substratem białko Hsp70 jest w stanie, za pośrednictwem swojej domeny NBD, oddziaływać z domeną M białka Hsp100. Prowadzi to do przekazania substratu związanego przez białko Hsp70 na białko Hsp100 i zniesienie jego inhibicji przez domenę M (Mogk, Kummer, i Bukau 2015).

Wykazano, że do aktywacji heksamery Hsp100 wymagane jest oddziaływanie z więcej niż jedną cząsteczką Hsp70 związaną z substratem. Sprawia to, że białka Hsp100 ulegają aktywacji w pobliżu agregatów białkowych, zazwyczaj wiązanych przez duże ilości systemów Hsp70, nie ulegają zaś aktywacji przez nowo syntetyzowane białka, zazwyczaj związane jedynie z pojedynczym systemem Hsp70 (Mogk, Kummer, i Bukau 2015; Liberek, Lewandowska, i Ziętkiewicz 2008).

Wysoce wydajna i specyficzna dezagregacja umożliwiona przez współpracę systemu Hsp70 z białkami Hsp100 jest wymagana dla utrzymania proteostazy w warunkach szoku cieplnego. Kluczową rolę zdaje się odgrywać fakt, że dezagregacja przez system Hsp70-Hsp100 umożliwia ponowne zwinięcie białek uwolnionych z agregatów do konformacji natywnej i tym samym przywrócenie ich fizjologicznej aktywności (reaktywacja). Pokazuje to eksperyment z wykorzystaniem zmodyfikowanego bakteryjnego białka Hsp100 (ClpB) – BAP. Wariant ten przekazuje rozwinięte substraty do proteazy ClpP, na skutek czego ulegają one degradacji, zamiast reaktywacji. Mimo iż wariant ten jest w stanie wydajnie współpracować z systemem Hsp70 w usuwaniu agregatów z komórek, nie jest on wystarczający do zapewnienia termotolerancji (Weibezahn i in. 2005).

W odróżnieniu od warunków szoku cieplnego, w warunkach fizjologicznych aktywność dezagregacyjna samego systemu Hsp70 zdaje się być wystarczająca dla utrzymania proteostazy. Może być to przyczyną tego, że białka Hsp100, występujące w komórkach bakterii, drożdży i roślin, nie są obecne w zwierzętach, z reguły rzadko narażonych na intensywny szok cieplny ze względu na wysoką mobilność i wydajne mechanizmy termoregulacji (Zolkiewski, Zhang, i Nagy 2012).

4.2 Małe białka szoku cieplnego

Małe białka szoku cieplnego (sHsp, z ang. *small Heat shock proteins*) są trzecią, obok dezagregaz Hsp100 i systemu Hsp70, grupą białek opiekuńczych zaangażowanych w proces dezagregacji. Ich nazwa wywodzi się od faktu, że charakteryzują się względnie niską masą cząsteczkową (12 – ok. 40 kDa). W odróżnieniu od pozostałych rodzin białek opiekuńczych, sHsp nie posiadają zdolności hydrolizy ATP (Haslbeck i Vierling 2015).

Stanowią one zróżnicowaną rodzinę białek, występującą we wszystkich grupach organizmów żywych, a także w niektórych bakteriofagach (De Jong, Leunissen, i Voorter 1993; Kriehuber i in. 2010; Bourrelle-Langlois i in. 2016). Różne grupy organizmów posiadają różną liczbę sHsp. Bardziej złożone organizmy, jak rośliny i zwierzęta, zazwyczaj posiadają wiele sHsp – genom człowieka koduje 10 różnych sHsp, genom nicienia *Caenorhabditis elegans* 18, a genom *Arabidopsis thaliana* – 19 (Gonçalves i in. 2021; Aevermann i Waters 2008; Siddique i in. 2008).

Proste, jednokomórkowe organizmy jak drożdże i bakterie zazwyczaj zawierają 1 (jak w przypadku *Vibrio harveyi*), 2 (np. *Escherichia coli*, *Saccharomyces cerevisiae*) lub 3 (np. *Salmonella typhimurium*) różne sHsp, chociaż od tej reguły zdarzają się wyjątki, jak w przypadku bakterii *Bradhyrhizobium japonicum*, zawierających przynajmniej 12 różnych sHsp (Obuchowski i in. 2019; Haslbeck i in. 2004; Tomoyasu i in. 2013; Münchbach, Nocker, i Narberhaus 1999).

W organizmach jednokomórkowych sHsp zdają się odgrywać rolę przede wszystkim w odporności na warunki stresowe, jak np. szok cieplny. Delecje genów kodujących sHsp w bakteriach i drożdżach w warunkach fizjologicznych mają niewielki efekt fenotypowy, powodują jednak zmniejszoną przeżywalność w warunkach intensywnego, długotrwałego szoku cieplnego lub obniżonej ekspresji białek opiekuńczych z systemu Hsp70 (Kuczyńska-Wiśnik i in. 2002; Mogk i in. 2003). sHsp zdają się również odgrywać rolę w utrzymywaniu komórek *Mycobacterium tuberculosis* w stanie długotrwałej fazy stacjonarnej, odgrywającym rolę w patogenezie wywoływanej przez nie gruźlicy (Yuan, Crane, i Barry 1996; Hu i in. 2006).

W organizmach zwierzęcych, znacznie mniej narażonych na szok cieplny, sHsp pełnią różnorodne funkcje związane przede wszystkim z ochroną białek przed chronicznym stresem. Wykazano ich rolę w kontroli proteostazy w procesie starzenia w *C. elegans* (Walker i Lithgow 2003), a zaburzenia ich funkcjonowania u ludzi mogą wiązać się z rozwojem neuropatii, miopatii czy chorób nowotworowych (Benndorf i in. 2014; Sun i MacRae 2005). α - Krystaliny w soczewce oka kręgowców chronią zawarte tam białka przed agregacją, zachowując przejrzystość soczewki i zapobiegając wystąpieniu zaćmy (Moreau i King 2012).

4.2.1 Struktura małych białek szoku cieplnego

Mimo dużej różnorodności sekwencji, sHsp charakteryzują się charakterystyczną, wysoce konserwowaną strukturą. Głównym jej elementem jest konserwowana domena α -krystalinowa (ACD, z ang. *Alpha – Crystallin Domain*), której nazwa wywodzi się od pierwszych opisanych sHsp – α A i α B krystalin, występujących w soczewce oka kręgowców. Domena ACD ma formę tak zwanej „kanapki β ” (ang. *beta sandwich*) – dwóch przeciwrównoległych β – kartek, z których jedna składa się z 4 wstążek β (β 2, β 3, β 8 i β 9), a druga z 3 wstążek β (β 4, β 5 i β 7) (Ingolia i Craig 1982; Clark i in. 2011; Hilario i in. 2011).

Wstążki β 4 i β 8, znajdujące się na krawędziach kartek tworzących domenę ACD, tworzą charakterystyczną kieszeń (Bepperling i in. 2012; Kennaway i in. 2005)(Ryc. 3B).

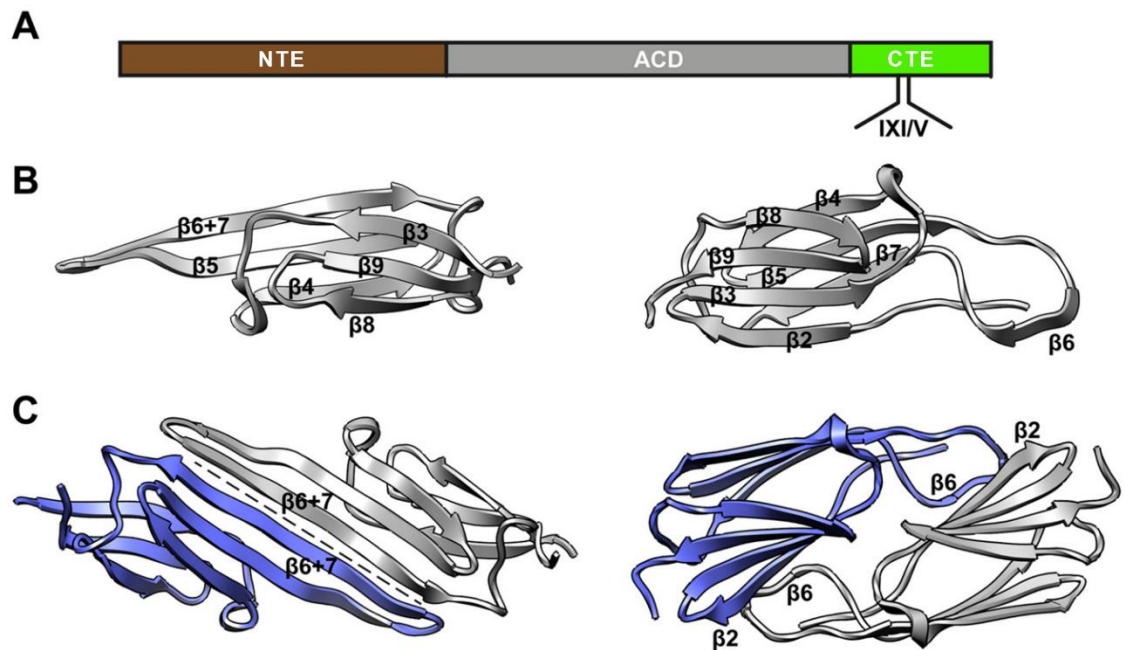
W bakteryjnych, archeobakteryjnych, roślinnych i drożdżowych sHsp wstążka β 6 domeny ACD znajduje się na charakterystycznej pętli, zwanej pętlą dimeryzacyjną. W zwierzęcych sHsp pętla ta nie występuje, a regiony homologiczne do wstążek β 6 i β 7 tworzą pojedynczą wstążkę β 6+7(Ryc. 3B) (Haslbeck, Weinkauff, i Buchner 2019).

Domena ACD flankowana jest przez dwa nieustrukturyzowane, nisko konserwowane regiony – NTE (ang. *N-Terminal Extension*), znajdujący się na końcu aminowym białka, oraz CTE (ang. *C-Terminal Extension*), stanowiący jego koniec karboksylowy. Niska zawartość struktur II – rzędowych w regionach NTE i CTE może być częściowo tłumaczona relatywnie wysoką zawartością prolin w ich sekwencji (Kriehuber i in. 2010). Region CTE zawiera wysoce konserwowany motyw IXI/V, zazwyczaj poprzedzony dodatkowo naładowanym aminokwasem (Haslbeck i Vierling 2015; Stróżecka i in. 2012)(Ryc. 3A).

Cechą charakterystyczną niemal wszystkich białek sHsp jest ich zdolność do tworzenia oligomerów, dynamicznie wymieniających podjednostki (Haslbeck i Vierling 2015; Stengel i in. 2010). Ich charakter różni się w zależności od konkretnego białka – do znanych struktur należą m.in. czworościenne 12 – mery tworzone przez białko Hsp 16.3 z *Mycobacterium tuberculosis* (Kennaway i in. 2005), 18 – mery o kształcie podwójnej piramidy i ośmiościenne 24-mery tworzone przez białko AgsA z *Salmonella typhimurium* (Mani i in. 2016), czy też oligomery fibrylarne tworzone przez białko IbpA z *Escherichia coli* (Ratajczak i in. 2010).

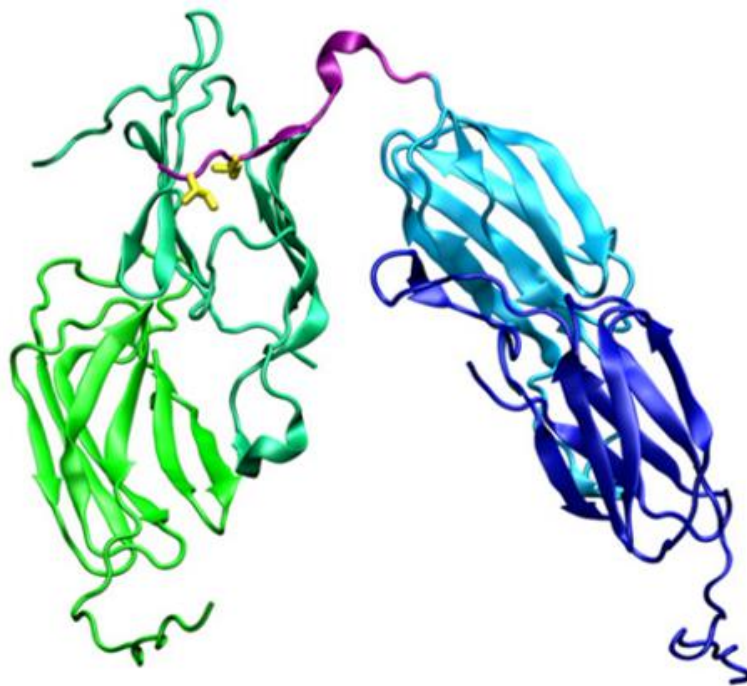
Podstawową jednostkę oligomerów sHsp stanowi dimer, tworzony na skutek oddziaływania między domenami ACD dwóch cząsteczek sHsp. W przypadku sHsp pochodzących z bakterii, archeonów, roślin i drożdży, dimeryzacja zachodzi przede wszystkim przez oddziaływanie między wstążką β 6, znajdującą się na tzw. pętli dimeryzacyjnej, ze wstążką β 2 na sąsiedniej cząsteczce. Dimery zwierzęcych sHsp,

niezawierających pętli dimeryzacyjnej, tworzone są przez oddziaływanie między wstążkami $\beta 6+7$ z dwóch różnych cząsteczek (Haslbeck, Weinkauff, i Buchner 2019)(Ryc. 3C).



Ryc. 3 Struktura sHsp: A) Schemat struktury sHsp pokazujący wzajemny układ domeny ACD i flankujących ją regionów NTE i CTE. W regionie CTE zaznaczono charakterystyczny, konserwowany motyw IXI/V. B) Struktura domeny ACD charakterystyczna dla zwierząt (Metazoa) (po lewej, na przykładzie ludzkiego HSPB5) oraz charakterystyczna dla bakterii, drożdży, archeobakterii i roślin (po prawej, na przykładzie Hsp17.7 z bakterii *D. radiodurans*); na modelach zaznaczono poszczególne wstążki β . C) Struktura dimeru tworzonego przez domeny ACD typowego dla sHsp z Metazoa (po lewej) i typowego dla sHsp z bakterii, drożdży, archeonów i roślin (po prawej). Zaadaptowano z (Haslbeck, Weinkauff, i Buchner 2019).

Organizacja dimerów sHsp w oligomery wyższego rzędu zachodzi przy udziale regionów NTE i CTE – domeny ACD pozbawione flankujących, nieustrukturyzowanych regionów są w stanie tworzyć jedynie formy dimeryczne (Stróżecka i in. 2012; Haslbeck, Weinkauff, i Buchner 2019). Najlepiej opisanym oddziaływaniem stabilizującym oligomery wyższego rzędu jest wiązanie konserwowanego motywu IXI/V, znajdującego się w regionie CTE, przez kieszeń $\beta 4$ - $\beta 8$ na domenie ACD sąsiedniego dimeru (Stróżecka i in. 2012; Haslbeck, Weinkauff, i Buchner 2019) (Ryc. 4). W przypadku zwierzęcych sHsp motyw ten może również oddziaływać z interfejsem dimerycznym tworzoną przez wstążki $\beta 6+7$ (Reinle, Mogk, i Bukau 2022).



Ryc. 4 Oddziaływanie między dwoma homodimerami białka sHsp IbpA z *E. coli* poprzez wiązanie regionu CTE przez kieszeń $\beta 4$ - $\beta 8$: Region CTE zaznaczono na fioletowo, a konserwowane izoleucyny tworzące motyw IXI/V zaznaczono na żółto. Zaadaptowane ze (Stróżecka i in. 2012)

Rola regionu NTE w organizacji dimerów sHsp w oligomery wyższego rzędu jest słabiej poznana, ze względu na trudności z uzyskaniem wysokiej rozdzielczości danych krystalograficznych dla tego regionu (Obuchowski, Karaś, i Liberek 2021). Najprawdopodobniej regiony NTE pochodzące z poszczególnych dimerów oddziałują ze sobą nawzajem, grupując się wewnątrz struktury oligomeru (Kennaway i in. 2005; Haslbeck i Vierling 2015). Regiony NTE niektórych sHsp zawierają motyw IXI/V, co sugeruje, że mogą one również oddziaływać z domenami ACD w sposób analogiczny do regionu CTE (Reinle, Mogk, i Bukau 2022).

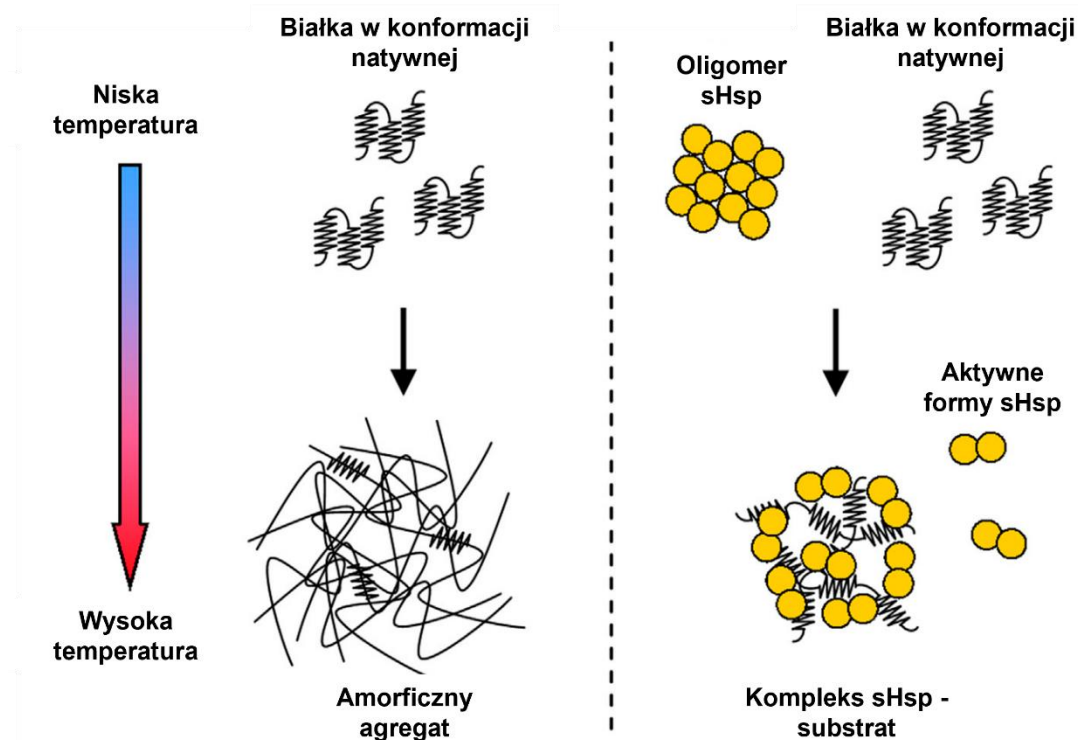
Oddziaływania stabilizujące oligomery wyższego rzędu są słabsze niż oddziaływania stabilizujące dimery sHsp. W połączeniu z elastycznym, nieustrukturyzowanym charakterem regionów NTE i CTE sprawia to, że oligomery sHsp mają charakter dynamiczny, a tworzące je jednostki dimeryczne podlegają nieustannej wymianie (Reinle, Mogk, i Bukau 2022).

Zmiany warunków zewnętrznych, zwłaszcza temperatury i pH, a także modyfikacje posttranslacyjne, jak np. fosforylacja, prowadzą do odwracalnych zmian stanu oligomerycznego sHsp, pełniących najprawdopodobniej funkcję regulatorową (Haslbeck i Vierling 2015).

W przypadku m.in. sHsp z *E. coli* wykazano, że duże oligomery mają charakter nieaktywny, a odwracalna deoligomeryzacja, zachodząca np. w warunkach szoku cieplnego, prowadzi do powstania aktywnych, mniejszych form (najprawdopodobniej dimerów) (Ratajczak, Zietkiewicz, i Liberek 2009).

4.2.2 Mechanizm działania i funkcja małych białek szoku cieplnego

Podstawową funkcją sHsp jest wiązanie i sekwestracja agregujących białek w kompleksach sHsp – substrat (ang. *sHsp – substrate assemblies*), w celu zapobieżenia ich dalszej, niekontrolowanej agregacji, do czasu, kiedy będą mogły ulec dezagregacji i przywróceniu do konformacji natywnej przez białka systemu Hsp70 i Hsp100. Aktywność tą określa się jako aktywność sekwestracji (Reinle, Mogk, i Bukau 2022)(Ryc. 5).



Ryc. 5 Wpływ sHsp na proces agregacji: sHsp są w stanie sekwestrować denaturowany, agregujący substrat w kompleksach zapobiegających jego dalszej agregacji. Kompleksy takie w warunkach *in vitro* charakteryzują się znacząco mniejszym rozmiarem niż amorficzne agregaty utworzone pod nieobecność sHsp, a substrat w nich związany jest sekwestrowany w konformacji zbliżonej do natywnej. Zaadaptowano z (Obuchowski, Karaś, i Liberek 2021).

Dokładny mechanizm wiązania substratów przez sHsp nie jest poznany, może on też różnić się w zależności od substratu i konkretnych białek sHsp. Kluczową rolę zdają się odgrywać w tym procesie regiony NTE i kieszeń $\beta 4$ - $\beta 8$ domeny ACD (Basha, Friedrich, i Vierling 2006; Jaya, Garcia, i Vierling 2009; Fuchs i in. 2009; Lee i in. 1997; Reinle, Mogk, i Bukau 2022). Fakt, że obszary te uczestniczą również w procesie

oligomeryzacji może tłumaczyć, w jaki sposób zmiany stanu oligomerycznego mogą regulować aktywność sHsp – rozpad oligomerów na mniejsze podjednostki odsłania miejsca wiązania substratu.

Kompleksy sHsp – substrat składają się z dwóch warstw. Centralny region kompleksu składa się z substratów białkowych i silnie związanych z nimi sHsp. Jest on otoczony warstwą sHsp, osłaniającą go przed oddziaływaniem z innymi nienatywnymi substratami i dalszą agregacją. Ta zewnętrzna warstwa ma charakter dynamiczny – tworzące ją sHsp są słabiej związane do substratu i ulegają ciągłej dysocjacji i ponownemu wiązaniu (Żwirowski i in. 2017).

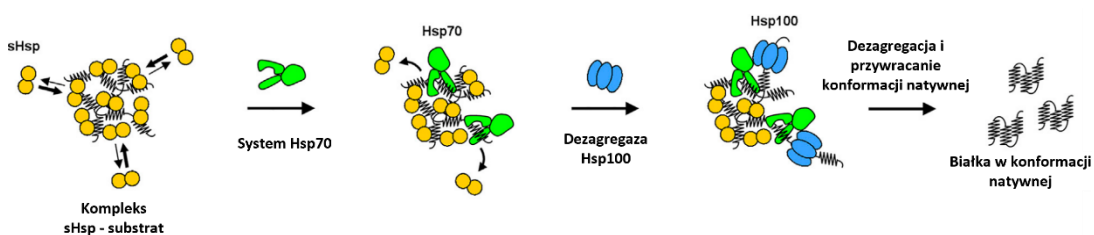
Wykazano, że dodatek peptydu o sekwencji regionu CTE w warunkach *in vitro* powoduje dysocjację sHsp tworzących tą warstwę zewnętrzną (Żwirowski i in. 2017). Sugeruje to, że oddziaływanie między regionem CTE, a kieszenią β 4- β 8 lub oddziaływanie między tą kieszenią a substratem odgrywa istotną rolę w stabilizacji tej warstwy.

Aby mogło dojść do utworzenia typowych kompleksów sHsp – substrat, białka sHsp muszą być obecne w momencie agregacji substratu, co umożliwia im wiązanie go na wczesnych etapach agregacji. sHsp dodane do wcześniej utworzonych agregatów są w stanie z nimi oddziaływać, potencjalnie tworząc zewnętrzną warstwę ochronną, nie są już jednak na tym etapie w stanie wpłynąć na rozmiar agregatu ani na konformację polipeptydów go tworzących (Stróżecka i in. 2012; Żwirowski i in. 2017).

Kompleksy sHsp z sekwestrowanymi białkami stanowią substrat dla procesu dezagregacji prowadzonego przez system Hsp70 i białka Hsp100. Białka sekwestrowane w tych kompleksach zazwyczaj znajdują się w konformacji zbliżonej do natywnej (Ungelenk i in. 2016). Co więcej, w doświadczeniach *in vitro* kompleksy takie posiadają znacząco mniejszy promień niż agregaty tworzące się pod nieobecność sHsp, a co za tym idzie wyższy stosunek powierzchni do objętości (Ratajczak, Zietkiewicz, i Liberek 2009). Sprawia to, że dezagregacja i przywracanie konformacji natywnej substratów znajdujących się w takich kompleksach przez system Hsp70 i białka Hsp100 może zachodzić z dużo większą wydajnością.

Co może wydawać się zaskakujące, wykazano, że w warunkach *in vitro* efektywność, z jaką sHsp stymulują dezagregację i przywrócenie konformacji natywnej substratów przez system Hsp70 i białka Hsp100 zależy w znacznym stopniu od stężenia białka Hsp70. W przypadku niskich stężeń systemu Hsp70 sHsp mogą wręcz powodować inhibicję zamiast stymulacji dezagregacji substratów. Podobnego efektu nie obserwuje się dla niskich stężeń białka Hsp100 (Żwirowski i in. 2017; Obuchowski i in. 2019).

Zjawisko to tłumaczy się faktem, że zewnętrzna warstwa sHsp, chroniąca substraty przed niekontrolowaną agregacją, osłania je również przed innymi białkami opiekuńczymi. Aby więc mógł zajść proces dezagregacji, system Hsp70 musi najpierw zastąpić sHsp tworzące zewnętrzną warstwę kompleksu (Ryc. 6). Zachodzi to na drodze prostej konkurencji o miejsca wiązania pomiędzy systemem Hsp70, a sHsp dynamicznie wiążącymi się do i dysocjującymi od kompleksu sHsp – substrat (Żwirowski i in. 2017; Obuchowski i in. 2019).



Ryc. 6 sHsp w procesie dezagregacji i przywracania substratów do konformacji natywnej: Zewnętrzna warstwa kompleksu sHsp – substrat, utrzymującego sekwestrowane białka w konformacji zbliżonej do natywnej, składa się z sHsp dynamicznie wiążących i dysocjujących od substratu, osłaniających go przed interakcjami potencjalnie prowadzącymi do niekontrolowanej agregacji. Białko Hsp70 konkuruje o miejsca wiązania na substracie z tymi sHsp, wypierając je z kompleksu. Hsp70 związane do substratu rekrutuje następnie dezagregazę Hsp100, co umożliwia dezagregację i przywrócenie substratu do konformacji natywnej. Zaadaptowano z (Obuchowski, Karaś, i Liberek 2021).

Wymusza to pewien kompromis pomiędzy aktywnością sekwestrazy białek sHsp i ich zdolnością do stymulacji procesu dezagregacji i przywracania konformacji natywnej substratów. sHsp silnie oddziałujące z substratami są w stanie efektywniej sekwestrować agregujące białka, jednak późniejsze przywrócenie konformacji natywnej związanym przez nie białkom wymaga wyższego poziomu systemu Hsp70. Z drugiej strony, białka sekwestrowane przez sHsp słabiej oddziałujące z substratami mogą być efektywnie przywrócone do konformacji natywnej przy niższym poziomie systemu Hsp70, jednak sam proces ich sekwestracji zachodzi dużo mniej wydajnie (Obuchowski i in. 2019).

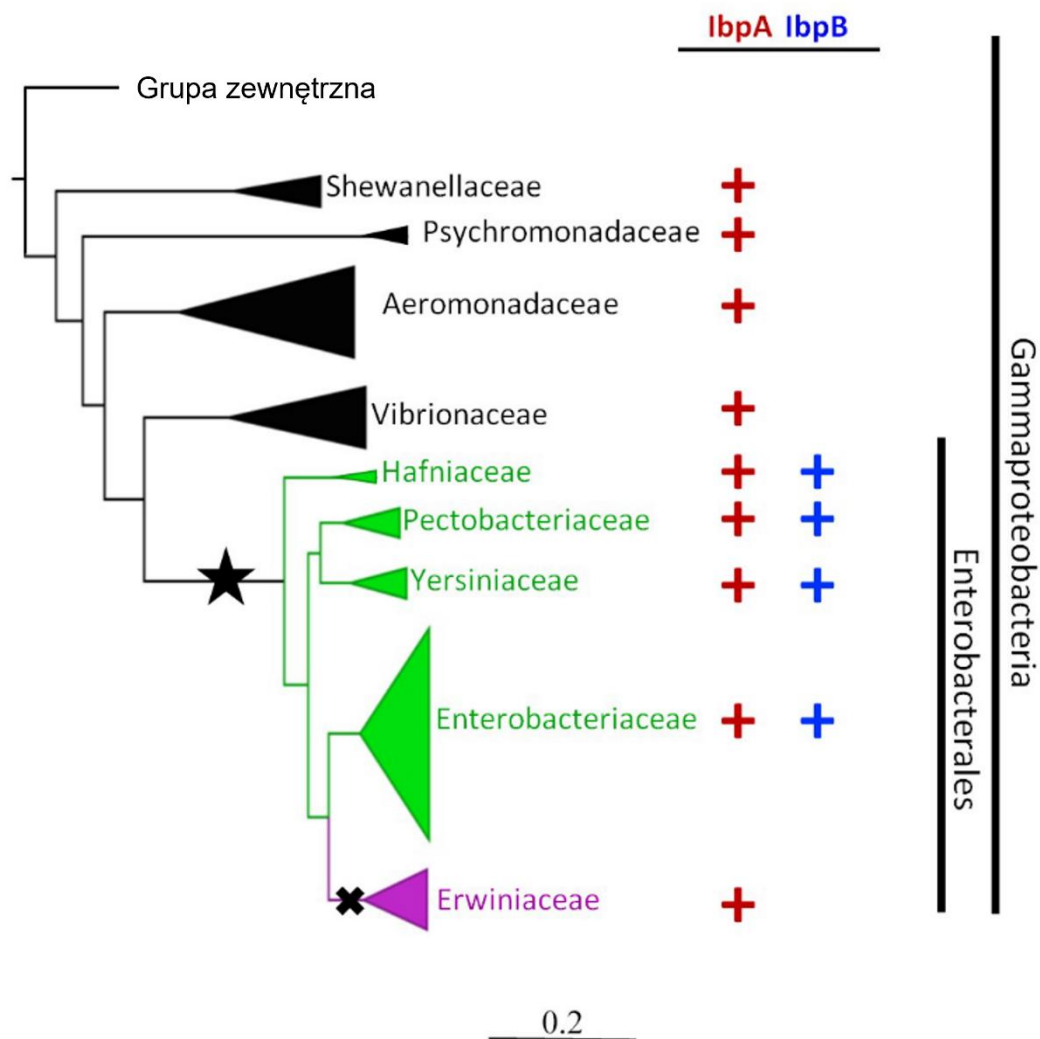
Postuluje się, że inhibicja dezagregacji przez sHsp przy niskich stężeniach systemu Hsp70 może odgrywać pewną rolę regulatorową, powstrzymując przywracanie konformacji natywnej sekwestrowanych substratów do momentu zakończenia warunków stresowych i ponownego zwiększenia dostępności systemu Hsp70 (Reinle, Mogk, i Bukau 2022).

4.3 Małe białka szoku cieplnego w *Enterobacterales*

Większość bakterii należących do rzędu *Enterobacterales* (jak *Hafniaceae*, *Pectobacteriaceae*, *Yersiniaceae* czy *Enterobacteriaceae*) posiada dwa sHsp – IbpA i

IbpB (z ang. *Inclusion body binding protein A*, *Inclusion body binding protein B*). Wyróżnia je to na tle pozostałych przedstawicieli klasy *Gammaproteobacteria*, posiadających jedynie pojedyncze sHsp – IbpA. Uważa się, że ten dwubiałkowy system powstał na skutek duplikacji ancestralnego genu *ibpA*, która zaszła w ostatnim wspólnym przodku rzędu *Enterobacterales* (Ryc. 7) (Obuchowski i in. 2019).

Wyjątek stanowi rodzina *Erwiniaceae*, która mimo przynależności do *Enterobacterales* posiada jedynie pojedyncze sHsp – IbpA. Rekonstrukcja historii ewolucyjnej pokazuje, że w rodzinie tej doszło do utraty genu *ibpB*, doprowadzając do powstania wtórnie pojedynczego sHsp (Ryc. 7) (Obuchowski i in. 2019).



- Pierwotnie pojedyncze IbpA
- System dwubiałkowy IbpA + IbpB
- Wtórnie pojedyncze IbpA
- ★ Duplikacja IbpA
- ✘ Utrata IbpB

Ryc. 7 Dystrybucja filogenetyczna sHsp w *Gammaproteobacteria*: *IbpA* obecne jest we wszystkich rodzinach należących do γ – proteobakterii. *IbpB*, powstałe na skutek duplikacji genu *ibpA* w ostatnim wspólnym przodku rzędu *Enterobacterales*, obecne jest w rodzinach *m.in.* *Hafniaceae*, *Pectobacteriaceae*, *Yersiniaceae* i *Enterobacteriaceae*. W rodzinie *Erwiniaceae* doszło do utraty genu *IbpB*. Rodziny zawierające pierwotnie pojedyncze *IbpA*, system dwubiałkowy składający się z białek *IbpA* i *IbpB* lub wtórnie pojedyncze *IbpA* (patrz wyżej) zaznaczono kolorami. Zaadaptowano z (Obuchowski i in. 2019).

4.3.1 Dwubiałkowy system sHsp w *Enterobacteriaceae*

Białka *IbpA* i *IbpB* z *E. coli* należącej do rodziny *Enterobacteriaceae*, będące przykładem dwubiałkowego systemu sHsp z *Enterobacterales*, są pierwszymi opisanymi bakteryjnymi sHsp i należą do najlepiej poznanych.

Białko *IbpA* z *E. coli* jest kanonicznym bakteryjnym sHsp, wykazującym właściwości podobne do ortologicznych *IbpA* z *Gammaproteobacteria*, w których nie doszło do duplikacji. Posiada ono wysokie powinowactwo do substratu i silną aktywność sekwestrazy, efektywnie wiążąc agregujące białka w kompleksach sHsp – substrat nawet gdy obecne jest w relatywnie niewielkim stężeniu. Ze względu na ową silną aktywność sekwestrazy, wydajna dezagregacja związanych przez nie substratów wymaga względnie wysokiego stężenia systemu Hsp70 (Obuchowski i in. 2019).

Białko *IbpB*, będące paralogiem *IbpA*, charakteryzuje się z kolei niskim powinowactwem do większości substratów i znikomą aktywnością sekwestrazy. Nie jest ono w stanie utworzyć kompleksów z większością substratów w warunkach *in vitro*. Ze względu na znikomą zdolność do samodzielnej sekwestracji agregujących substratów, białko *IbpB* nie jest w stanie samodzielnie stymulować ich dezagregacji przez system Hsp70 (Ratajczak, Zietkiewicz, i Liberek 2009; Obuchowski i in. 2019).

Oba sHsp (*IbpA* i *IbpB*) są w stanie ściśle ze sobą współpracować - gdy w układzie obecne jest zarówno *IbpA* jak i *IbpB*, funkcjonują one preferencyjnie jako heterodimer (Piróg i in. 2021). Współpraca ta umożliwia białku *IbpB* wiązanie substratów za pośrednictwem białka *IbpA*, dzięki czemu może ono wchodzić w skład kompleksów sHsp – substrat. Obecność *IbpB* ułatwia dysocjację *IbpA* związanych do substratu w czasie dezagregacji, zmniejszając stężenie systemu Hsp70 wymagane, aby sHsp mogły wydajnie stymulować dezagregację (Obuchowski i in. 2019).

W rezultacie system współpracujących ze sobą białek *IbpA* i *IbpB* jest w stanie zarówno efektywnie sekwestrować agregujące substraty jak i skutecznie stymulować ich dezagregację przez system Hsp70 i białka Hsp100, nawet przy względnie niskim stężeniu systemu Hsp70 (Obuchowski i in. 2019). Obecność białka *IbpB* umożliwia tym samym bakteriom z rzędu *Enterobacterales* częściowe ominięcie potrzeby kompromisu między aktywnością sekwestrazy sHsp, a ich zdolnością do stymulacji dezagregacji i ponownego zwijania białek do konformacji natywnej.

4.3.2 Wtórnie pojedyncze IbpA w *Erwiniaceae*

W odróżnieniu od pozostałych przedstawicieli rzędu *Enterobacterales* (w tym rodziny *Enterobacteriaceae*), bakterie należące do rodziny *Erwiniaceae* (składającej się m. in. z rodzajów *Erwinia*, *Pantoea*, *Mixta*, *Tatumella* i *Phaseliobacter*) posiadają jedynie pojedyncze sHsp - IbpA. Analiza historii ewolucyjnej sHsp w *Gammaproteobacteria* wykazuje, że ostatni wspólny przodek *Erwiniaceae* i ich kladu siostrzanego *Enterobacteriaceae* najprawdopodobniej posiadał typowy dla *Enterobacterales* system dwubiałkowy, a w rodzinie *Erwiniaceae* doszło do utraty genu *ibpB*. Z tego względu, dla odróżnienia od IbpA występującego w organizmach, które w swojej historii ewolucyjnej nigdy nie posiadały systemu dwubiałkowego (tzw. pierwotnie pojedynczego IbpA), białko IbpA z rodziny *Erwiniaceae* określa się jako wtórnie pojedyncze (ang. *secondarily single*) białko IbpA (Obuchowski i in. 2019).

Wtórnie pojedyncze IbpA z *Erwiniaceae*, wbrew temu, czego można by się było spodziewać, nie charakteryzuje się silną aktywnością sekwestrazy typową dla IbpA z systemu dwubiałkowego czy też pierwotnie pojedynczych IbpA występujących w *Gammaproteobacteria* innych niż *Enterobacterales*. Jego zdolność do wiązania i sekwestracji agregujących substratów jest znacząco niższa niż IbpA wchodzącego w skład systemu dwubiałkowego czy pierwotnie pojedynczych IbpA. Jest ona jednak wyższa niż w przypadku białek IbpB, tak że wtórnie pojedyncze białka IbpA są wciąż w stanie samodzielnie sekwestrować agregujące substraty i stanowią funkcjonalne sHsp. Co więcej, ich zdolność do stymulacji dezagregacji i ponownego zwijania sekwestrowanych substratów przez system Hsp70 jest wyższa niż w przypadku pozostałych pojedynczych IbpA, bardziej zbliżona do współpracujących ze sobą białek IbpA i IbpB z systemu dwubiałkowego (Obuchowski i in. 2019). Wskazuje to, że w rodzinie *Erwiniaceae* doszło do uproszczenia dwubiałkowego systemu sHsp do systemu złożonego jedynie z pojedynczego białka IbpA, przy zachowaniu w znacznym stopniu jego zwiększonej zdolności do stymulacji dezagregacji i ponownego zwijania substratów przy niższym stężeniu systemu Hsp70.

Molekularne podstawy różnic w aktywności wtórnie pojedynczego IbpA z *Erwiniaceae* i IbpA z siostrzanego kladu *Enterobacteriaceae*, które stanowi część systemu dwubiałkowego, nie są wciąż dobrze poznane. Zrozumienie, w jaki sposób IbpA z *Erwiniaceae* uniezależniło się od współpracy z IbpB mogłoby potencjalnie rzucić światło na mechanizmy oddziaływania bakteryjnych sHsp z substratami i stabilizacji zewnętrznej warstwy kompleksów sHsp - substrat.

5. Cel Pracy

Celem niniejszej pracy jest zbadanie ewolucji uproszczonego systemu małych białek szoku cieplnego w *Erwiniaceae* poprzez rekonstrukcję historii ewolucyjnej białek IbpA w *Enterobacterales*, ze szczególnym uwzględnieniem rodziny *Erwiniaceae*, identyfikację zmian sekwencji aminokwasowej kluczowych dla nabycia przez wtórnie pojedyncze IbpA w *Erwiniaceae* nowej funkcjonalności oraz zaproponowanie roli tych zmian na poziomie molekularnym.

6. Materiały i Metody

6.1 Materiały

6.1.1 Szczepy bakterii

- ***E. coli* BL21(DE3)** – nadprodukcja białek; genotyp: *fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS* (Zasoby własne Pracowni Biochemii Białek)
- ***E. coli* Mach1™** – klonowanie genów i długoterminowe przechowywanie plazmidów; genotyp: *F- φ80(lacZ)ΔM15 ΔlacX74 hsdR(rK- mK+) ΔrecA1398 endA1 tonA* (Thermo Scientific)
- ***E. coli* BB6406** – uzyskanie lizatu służącego jako jeden z substratów w analizie oddziaływania sHsp z agregatami białkowymi; genotyp: *araD139 Δ(argF-lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR PA1/lacO-1 dnaK, dnaJ, lacIq, ΔibpAB* (Mogk i in. 2003)

6.1.2 Plazmidy

- **pET3a** – w celu nadprodukcji białek IbpA_{E. amylovora}, AncA₀ i jego mutantów, AncA₁, AncA₁ H66Q D109G, AncA₀Alt, AncA₁Alt, IbpA_{E. amylovora} H67Q D110G i IbpA_{E. coli} Q66H G109D; AmpR, promotor T7 (GenScript)
- **pET28a** – w celu nadprodukcji peptydu C ze znacznikiem His₆-SUMO i domen ACD; KanR, promotor T7 (GenScript)
- **pET28b** – w celu nadprodukcji białka IbpB_{E. coli}; KanR, promotor T7 (GenScript)

6.1.3 Startery oligonukleotydowe

Nazwa	Sekwencja	Zastosowanie
T50A_rev T50A_fwd H66Q_rev H66Q_fwd R73K_fwd R73K_rev E78G_fwd E78G_rev D109G_rev D109G_fwd R111N_rev R111N_fwd L137N_fwd L137N_rev	GCAATGGCGATGCGGTAGTGGTTTTTC GCATCACCGTTGCTGTGGC ATCCTGCGCCGTGATGTCCA TCACGGCGCAGGATAATCTGTAA ATTGTTAAAGGTGCGCACGC GCACCTTTAACAATAACAGATTATCATGCGCCG CACGCAGGTGAACAGCCGGAGC TGTTACCTGCGTGCGCAC GCGCACCACGAACATGGATGTG TCGTGGTGCGCGTCTCG CGAGATTCGCATCACGAACATGGATG GATGCGAATCTCGAGAACGGCC AAATCAACTAAGGATCCGGCTGCTAAC CCTTAGTTGATTTCAATACGACGCGG	Introdukcja mutacji do białka AncA ₀₊₇
Q66H_fwd Q66H_rev G109D_fwd G109D_rev	GCGCATGACAATCTGCTTATCGTG TTGTCATGCGCGGTGATGTC CATGTTGCTGATGCTAATTTGGAGAAC TTAGCATCACGAACATGAATGTGC	Introdukcja mutacji do białka AncA ₀
E1H66Q_rev E1H66Q_fwd E1D109G_rev E1D109G_fwd	GGTTATCCTGGCTGGTGATGTCCAG CCAGGATAACCTGCTTATCGTTTCG CGAGCACC GCGCACGTGGATG GCGCGGTGCTCGTCTGGAGAAC	Introdukcja mutacji do białka AncA ₁

Erwinia_H67Q_fwd Erwinia_H67Q_rev Erwinia_D110G_fwd Erwinia_D110G_rev	TCCCAGGACAACCTGCTTTCTG GCAGGTTGTCCTGGGAGGTGATCTC GTTTCGTGGTGCGGTCTGGAG CGCACCACGAACATGCACGTG	Introdukcja mutacji do białka <i>lbpA_{E. amylovora}</i>
Escherichia_G109D_fwd Escherichia_G109D_rev Escherichia_Q66H_fwd Escherichia_Q66H_rev	CGTGATGCTAATTTGGTTAATGGCCTG CCAAATTAGCATCACGAACATGAATGTT ACCGCTCATGATAATTTGTTGGTTG TATCATGAGCGGTGATCTCAAG	Introdukcja mutacji do białka <i>lbpA_{E. coli}</i>

Startery oligonukleotydowe zostały zsyntetyzowane *de novo* przez firmę Genomed

6.1.4 Syntetyczne geny

- ***ancA₀***: ATGAGGAATTTTGATCTATCACCTTATATAGAAGCGCGATCGGCTTTG
ATCGTCTGTTCAACCTGTTGGAGTCTAACCGAATCAAAGCAACGGCGGTTATC
CACCGTATAACGTTGAACTGGTGGATGAAAACCACTATCGTATCGCCATTGCTG
TGGCGGGCTTCGCCGAGTCCGAATTGGACATCACCGCGCAGGACAATCTGCTTA
TCGTGAAAGGTGCGCATGCAGGTGAACAGCCGGAGCGCACGTACCTGTACCAAG
GTATTGCGGAACGTAATTTTGAGCGCAAGTTCCTCAACTGGCGGAGCACATTCATG
TTCGTGGTGCTAATTTGGAGAACGGCCTGCTCTACATCGACCTGGAGCGCGTCCG
TTCCGGAAGCAATGAAACCGCGTCGTATTGAAATCAACTAA
- ***ancA₁***: ATGAGGAATTTTGATCTATCACCTTATATAGAAGCGCAATTGGCTTTG
ATCGTCTGTTTAACTTGTGGAGAGCAATCAAACCAATCCAATGGCGGTTATC
CACCGTATAACGTGGAATTGGTTGATGAAAACCACTACCGCATCACGATCGCCG
TTGCGGGTTTCGCCAGTCTGAACTGGACATCACCGCCATGATAACCTGCTTA
TCGTTTCGTGGTGACATGCGGAGGAACAGCCGGAGCGCACCTATCTGTACCAGG
GTATTGCGGAGCGTAATTTTGAACGTAAGTTCCTCAACTGGCGGAGCACATCCACG
TGCGCGACGCTCGTCTGGAGAACGGCCTGCTCTACATCGACCTGGAGCGCGTCCG
TGCCGGAAGCTATGAAACCGCGTCGTATTGAAATCTGAAATAA
- ***ancA₀₊₇***: ATGAGAAATTTTGATCTATCACCTTATATAGGTCTGCGATCGGCTTC
GATCGTTTGTTHAACTTGTGGAGTCCAATCAAACCAAAGCAATGGCGGTTAT
CCACCGTATAACGTGGAACTGGTGGACGAAAACCACTACCGCATCACCATTTGCT
GTGGCCGGTTTCGCCGAGAGCGAACTGGACATCACGGCGCATGATAATCTGTTA
ATTGTTTCGTGGTGCGCACGCAGAGGAACAGCCGGAGCGCACCTATCTGTACCAG
GGTATTGCTGAGCGCAACTTTGAACGCAAGTTCAGCTGGCGGAGCACATCCAT
GTTCGTGATGCGCGTCTCGAGAACGGCCTGCTGTACATCGACCTGGAGCGTGT
GTCCCGGAAGCAATGAAACCGCGTCGTATTGAAATCCTGTAA
- ***ancA₀ ACD***: ATGGGAAGTTCACATCACACCACCACCATTCTTCCGGCCTGGTCC
CGCGTGGCAGCCATGGTGGTTATCCGCCATACAATGTTGAACTCGTGGATGAAA
ACCACTATCGTATTGCGATTGCTGTGGCGGGTTTTGCGGAGAGCGAACTGGACA
TCACCGCACAGGACAACCTGCTTATCGTGAAAGGTGCGCATGCTGGCGAACAAC
CGGAACGCACGTATCTGTACCAAGGTATTGCCGAGCGTAATTTTCGAGCGCAAGT
TCCAGCTGGCAGAGCACATCCACGTTTCGTGGCGCGAACCTGGAGAACGGCTTGT
TGTACATCGATCTGGAGAGATAA
- ***ancA₀ Q66H G109D ACD***: ATGGGATCTTACATCATCACACCACCACAGTTCCG
GCCTGGTTCCGCGTGGTAGCCACGGCGGTTATCCACCGTATAACGTGGAATTGG
TCGATGAGAATCATTATCGTATTGCGATCGCAGTTGCGGGTTTTGCCGAGAGCG

AACTGGACATCACCGCACACGACAACCTGCTGATTGTTAAAGGTGCGCACGCTG
GCGAACAGCCGGAACGTACGTACCTGTACCAAGGTATTGCGGAGCGCAACTTCG
AGAGAAAGTTCCAGCTGGCGGAACATATCCACGTGCGTGATGCTAATCTCGAGA
ACGGCTTGTATACATCGACCTGGAGCGCTAA

- **his₆ - sumo -peptyd C:** ATGCATCATCATCATCACGGCAGCGGCCTGGTGCCG
CGCGGCAGCGCTAGCATGTTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAG
GTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCGATGGA
TCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATG
GAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAGATTCTTGTAC
GACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGAT
AACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGTGGAGGGTCAGGTGGA
CCCGAAGCTATGAAACCGCGCCGTATTGAGATCAACTAA

Geny zostały zsyntetyzowane *de novo* przez firmę GenScript na podstawie rekonstruowanych sekwencji aminokwasowych i wstawione do plazmidu pET3a (w przypadku *ancA₀*, *ancA₁* i *ancA₀₊₇*) lub pET28a (w przypadku genów kodujących domeny ACD i peptyd C).

6.1.5 Sekwencje aminokwasowe wykorzystane do analizy filogenetycznej i rekonstrukcji ancestralnej:

- **Numery akcesji sekwencji aminokwasowych pozyskanych z bazy danych UniProt:**

C5BHD3, A0A097QX83, A0A150T1G1, W8TUF5, W0LEB5, L0MCB7, A8G7R8, E0SEQ3, Q6CYV2, B2VCD6, E5BAR7, A0A0A1B7C9, A0A0A0Z1A4, E6WHH2, D4GFY5, E1SJI2, A0A089UAZ7, I2B3P3, A0A0G3QF10, A6TFZ0, A0A0H3FP83, A0A089PH96, A0A1S2BE20, L0LWV8, X2KC60, D2THX5, A0A0K0HFM0, A9MJW8, A8ACK0, A4W4S3, POC054

- **Numery akcesji sekwencji aminokwasowych pozyskanych z bazy danych NCBI:**

WP_023656845.1, ADP10815.1, WP_123334577.1, WP_212409578.1,
WP_062744073.1, WP_016192433.1, WP_034889306.1, WP_124232393.1,
WP_067700519.1, WP_147193638.1, WP_125289205.1, WP_152320709.1,
WP_158782426.1, WP_052901051.1, WP_024966304.1, WP_078002538.1,
WP_017800319.1, WP_092879054.1, WP_103060270.1, WP_085069121.1,
WP_104955624.1, WP_193406688.1, WP_160622854.1, HBZ16520.1,
WP_215848074.1, WP_081140557.1, WP_208443522.1, WP_194761869.1,
WP_145885986.1, WP_007893512.1, HCX00537.1, WP_031591032.1,
WP_172913532.1, WP_123802177.1, WP_130832981.1, WP_048785291.1,
MXP49787.1, MBA2817917.1, WP_176973343.1, WP_192236804.1,
WP_159932146.1, WP_029685321.1, WP_087486742.1, WP_038016416.1,
WP_028685261.1

6.1.6 Sekwencje nukleotydowe wykorzystane do analizy selekcji naturalnej:

- **Numery akcesji sekwencji nukleotydowych pozyskanych z bazy NCBI:**

NZ_CP084521.1: 5330-5743, CP009706.1: 172869-173279, CP045258.1: 12657-13067, CP007448.2: 18060-18470, CP007044.2: 899182 - 899592, CP003942.1: 12273-12680, CP038467.1: 3415460-3415870, CP002038.1: 100743-101153, BX950851.1: 4942277-4942687, CU468135.1: 3841987-3842403, FR719198.1: 304583-304999, CP009866.1: 3808609-3809022, CP009880.2: 1221095-1221508, CP002433.1: 13738-14151, CP054912.1: 2997265-2997747, CP002206.1: 3489011-3489424, CP009458.1: 957142-957552, LR133996.1: 10858-11268, CP011602.1: 4624766-4625176, CP090222.1: 5265032-5265487, CP050067.1: 16620-17030, CP009450.1: 854887- 855297, NZ_BCZS01000026.1: 34064-34474, CP003938.1: 14328-14738, CP030029.1: 34627-35037, CP038008.1: 3488060-3488470, CP045766.1: 1032344-1032754, CP053401.1: 4125750-4126166, CP052059.1: 32748-33158, CP000653.1: 14693-15100, CP119120.1: 4786975-4787385, NZ_CAHS01000023.1: 93442-93858, CP002124.1: 1240593-1241086, NZ_RJVB01000005.1: 373145-372726, CP046582.1: 16769-17185, CP022725.1: 4002737-4003156, CP013970.1: 3590047-3590460, NZ_JRUQ01000019.1: 223897-223481, NZ_RQSB01000009.1: c77611-77195, NZ_CDPK01000001.1: c279796-279380, CP034363.1: 9240-9653, NZ_RHUM01000007.1: 69812-70228, NZ_RQVV01000015.1: 79119-79535, NZ_JRXE01000025.1: c10120-9704, NZ_JFGT01000004.1: 310317-310736, NZ_MRUL01000005.1: c146170-145754, NZ_JAGGMQ010000001.1: c2363110-2362691, NZ_FOVC01000010.1: 100737-101153, NZ_NWUO01000009.1: c119698-119282, NZ_CP019706.1: c1591751-1591335, NZ_CP026377.1: 59987-60403, NZ_JACFX010000017.1: 38919-39335, NZ_CP028271.1: 3788769-3789182, JAPWKD010000022.1: c244063-243647, NZ_MWUE01000023.1: 59723-60139, NZ_CAJOSF010000008.1: 173282-173698, NZ_JACAFO010000008.1: 147241-147657, NZ_CP032702.1: 9218-9634, NZ_AKIU01000090.1: c79202-78789, DPPD01000429.1: 4536-4952, NZ_OW970315.1: 9841-10257, NZ_JADWWL010000014.1: 140326-140742, NZ_RMVG01000015.1: 59041-59457, NZ_JUQR01000264.1: c20649-20233, SZZY01000007.1: 63197-63613, SMDS01000023.1: 61674-62090, NZ_JACAPA010000052.1: 66934-67347, NZ_JACVBT010000006.1: 62907-63320, NZ_SZZW01000007.1: c20608-20192, NZ_ATMI01000011.1: 13154-13576, NZ_CP015581.1: 12390-12815, NZ_JPKR02000005.1: 814901-815326

6.1.7 Białka

- **DnaK** – zasoby własne Pracowni Biochemii Białek
- **DnaJ** – zasoby własne Pracowni Biochemii Białek
- **GrpE** – zasoby własne Pracowni Biochemii Białek
- **ClpB** – zasoby własne Pracowni Biochemii Białek
- **BSA** – Albumin standard (Thermo Scientific)

- **Lucyferaza** – QuantiLum® Recombinant Luciferase (Promega)
- **Kinaza kreatyny** – kinaza kreatyny z mięśnia króliczego (Sigma Aldrich)
- **Trombina** – Thrombin from bovine plasma (SERVA)
- **DpnI** (Thermo Scientific)
- **Polimeraza DNA PfuUltra** – PfuUltra II Fusion HS DNA Polymerase (Agilent)
- **IbpA_{E.coli}** – zasoby własne Pracowni Biochemii Białek
- **IbpB_{E.coli}** – oczyszczone w ramach tej pracy
- **IbpA_{E.amylovora}** - oczyszczone w ramach tej pracy
- **IbpA_{E.coli} Q66H G109D** - oczyszczone w ramach tej pracy
- **IbpA_{E.amylovora} H67Q D110G** - oczyszczone w ramach tej pracy
- **AncA₀ i pochodne (AncA₀ Q66H, AncA₀ G109D, AncA₀ Q66H G109D)** – oczyszczone w ramach tej pracy
- **AncA₁ i AncA₁ H66Q D109G** – oczyszczone w ramach tej pracy
- **AncA₀Alt, AncA₁Alt** – oczyszczone w ramach tej pracy
- **AncA₀+7 i pochodne (AncA₀+7 T50A, AncA₀+7 H66Q, AncA₀+7 R73K, AncA₀+7 E78G, AncA₀+7 D109G, AncA₀+7 R111N, AncA₀+7 L137N)** – oczyszczone w ramach tej pracy
- **Domeny ACD białek AncA₀ i AncA₀ Q66H G109D** – oczyszczone w ramach tej pracy
- **Domena ACD białka IbpA_{E. coli}** – zasoby własne Pracowni Biochemii Białek
- **His₆ – SUMO** – oczyszczone w ramach tej pracy
- **His₆ – SUMO – peptyd C** - peptyd C (region CTE białka AncA₀ - PEAMKPPRIEIN) z dołączonym znacznikiem His₆ – SUMO; oczyszczony w ramach tej pracy

6.1.8 Pożywki

- **LB** – 1% Bacto™ peptone, 1% NaCl, 0,5% ekstrakt drożdżowy
- **LA** - 1% Bacto™ peptone, 1% NaCl, 0,5% ekstrakt drożdżowy, 1,5% agar

6.1.9 Antybiotyki

- **Ampicylina** (ROTH)
- **Kanamycyna** (BioShop)
- **Chloramfenikol** (BioShop)

6.1.10 Bufory

- **bufor do lizy A** - 50 mM Tris pH 7,5, 50 mM NaCl, 5mM EDTA, 10% glicerol 5 mM β-merkaptioetanol
- **bufor A** - 40 mM Tris pH 7,5, 50 mM NaCl, 10% glicerol, 5 mM β – merkaptioethanol, 6M mocznik
- **bufor B** - 40 mM Tris pH 8,5, 50 mM NaCl, 10% glicerol, 5 mM β – merkaptioetanol
- **bufor FL** - 50 mM Tris pH 7,5, 150 mM KCl, 10 % glicerol, 5 mM β – merkaptioetanol

- **bufor do lizy B** - 40 mM Tris pH 7,4, 50 mM NaCl, 5mM EDTA, 10% glicerol 5 mM β -merkaptobetanol
- **bufor C** - 40 mM Tris pH 7,5, 25 mM NaCl, 10% glicerol, 5 mM β – merkaptobetanol, 6M mocznik
- **buforu D** - 40 mM Tris pH 7,4, 500 mM NaCl, 10 % glicerol, 5 mM β – merkaptobetanol, 6M mocznik
- **bufor E** - 40 mM Tris pH 8,5, 150 mM NaCl, 10 % glicerol, 5 mM β – merkaptobetanol
- **bufor G** - 40 mM Tris pH 7,4, 500 mM NaCl, 10% glicerol, 5 mM β -merkaptobetanol, 6M mocznik
- **bufor H** - 40 mM Tris pH 7,4, 500 mM NaCl, 10% glicerol, 5 mM β -merkaptobetanol, 360 mM imidazol, 6M mocznik
- **bufor I** - 50 mM Tris pH 8,5, 150 mM KCl, 10% glicerol, 5 mM β -merkaptobetanol, 1 mM CaCl_2
- **bufor J** - 50 mM Tris pH 8,5, 150 mM KCl, 10% glicerol, 5 mM β -merkaptobetanol, 6M mocznik
- **bufor FB'** - 50 mM Tris pH 8,5, 150 mM KCl, 5 mM β – merkaptobetanol
- **bufor FB** - 50 mM Tris pH 7,5, 150 mM KCl, 5 mM β – merkaptobetanol
- **bufor K** - 40mM Tris pH 7,5, 100 mM NaCl, 10% glicerol, 5 mM β – merkaptobetanol, 10 mM imidazol
- **bufor L** - 40 mM Tris pH 7,5, 100 mM NaCl, 10% glicerol, 5 mM β -merkaptobetanol, 40 mM imidazole
- **bufor M** - 40 mM Tris pH 7,5, 100 mM NaCl, 10% glicerol, 5 mM β -merkaptobetanol, 250 mM imidazol
- **bufor F** - 150 mM KCl, 50 mM Tris pH 7,5, 20 mM MgCl_2 , 2.5 mM DTT

6.2 Metody

6.2.1 Analiza filogenetyczna i rekonstrukcja ancestralna

77 sekwencji aminokwasowych ortologów IbpA z *Enterobacterales* zostało pozyskanych z baz UniProt i NCBI, a następnie zestawionych przy użyciu programu Clustal Omega (Sievers i in. 2011). Model ewolucji JTT + R3 został wyznaczony przy pomocy programu iq-tree jako najlepiej opisujący uzyskane zestawienie, na podstawie kryterium BIC (ang. *Bayesian Information Criterion*) (Kalyaanamoorthy i in. 2017; Nguyen i in. 2015). Drzewo filogenetyczne najlepiej opisujące pokrewieństwo analizowanych sekwencji zostało następnie wyznaczone przy pomocy programu iq-tree na podstawie 328 powtórzeń analizy maksymalnej wiarygodności (ang. *Maximum Likelihood*) ze 100 powtórzeniami analizy typu bootstrap (Nguyen i in. 2015).

Rekonstrukcja sekwencji ancestralnych została przeprowadzona na podstawie drzewa filogenetycznego i zestawienia sekwencji uzyskanych na poprzednim etapie analizy. Przy użyciu programu FastML, opartego na algorytmie maksymalnej

wiarygodności (ang. *maximum likelihood*), została przeprowadzona rekonstrukcja marginalna (ang. *marginal reconstruction*) najbardziej wiarygodnych sekwencji ancestralnych w oparciu o model ewolucji JTT z parametrem gamma z optymalizacją długości gałęzi drzewa filogenetycznego (Ashkenazy i in. 2012; Cohen i Pupko 2011; Cohen i in. 2008; Jones, Taylor, i Thornton 1992; Pupko i in. 2002; Simmons i Ochoterena 2000).

Alternatywne warianty sekwencji ancestralnych zostały uzyskane przez substytucję najbardziej wiarygodnego aminokwasu na niepewnych pozycjach (definiowanych jako pozycje, na których prawdopodobieństwo *a posteriori* wystąpienia więcej niż jednego aminokwasu było większe lub równe 0,2) na aminokwas o drugim największym prawdopodobieństwie *a posteriori* (Eick i in. 2016).

6.2.2 Synteza genów *de novo*

Geny kodujące rekonstruowane sekwencje aminokwasowe (zarówno warianty najbardziej wiarygodne, jak i alternatywne), oraz gen *ancA₀₊₇* zostały zsyntetyzowane *de novo* i wklonowane do plamidu pET3a przez firmę zewnętrzną (GenScript), przy założeniu użycia kodonów zoptymalizowanego do ekspresji w *E. coli*.

Geny kodujące konstrukt His₆ – SUMO – peptyd C, jak i domeny ACD białek AncA₀ i AncA₀ Q66H G109D zostały zsyntetyzowane *de novo* i wklonowane do plamidu pET28a przez firmę zewnętrzną (GenScript), przy założeniu użycia kodonów zoptymalizowanego do ekspresji w *E. coli*.

6.2.3 Mutagenеза miejscowo - specyficzna

Mutacje punktowe zostały wprowadzone do genów *ancA₀*, *ancA₁*, *ibpA_{E.coli}*, *ibpA_{E.amylovora}* i *ancA₀₊₇* za pomocą mutagenезы miejscowo – specyficznej z wykorzystaniem reakcji PCR ze starterami zawierającymi fragmenty niekomplementarne.

Termofilna polimeraza DNA PfuUltraII została dodana do uprzednio rozgrzanej do 95 °C mieszaniny reakcyjnej (tzw. metoda gorącego startu – ang. *hot start*), zawierającej plazmid matrycowy, mieszaninę trifosforanów deoksynukleotydów oraz startery oligonukleotydowe w buforze dedykowanym dla polimerazy PfuUltra. Stężenia składników reakcji dobrano zgodnie z protokołem producenta polimerazy. Reakcję PCR prowadzono przez 19 cykli:

1. Denaturacja DNA w 95 °C przez 20 s.
2. Przyłączanie starterów do matrycy DNA w 56 °C przez 20s.
3. Elongacja starterów przez polimerazę DNA w 68 °C przez 1 min 16s.

Po zakończeniu 19 cykli reakcji PCR przeprowadzono dodatkowy etap elongacji w 68 °C przez 3 min.

Matrycowe DNA zostało następnie usunięte przez trawienie enzymem restrykcyjnym DpnI przez 1h w 37 °C. Mieszanina reakcyjna była następnie inkubowana przez 20 min w 80 °C w celu inaktywacji DpnI. Długość produktów reakcji PCR została zweryfikowana przy użyciu elektroforezy agarozowej.

Produktami PCR transformowano następnie komórki *E. coli* Mach1, które następnie wysiano na podłożu LA z dodatkiem ampicyliny w celu selekcji komórek transformowanych.

Po przeprowadzeniu mutagenyzy sekwencję uzyskanych plazmidów zweryfikowano przez sekwencjonowanie.

6.2.4 Izolacja plazmidowego DNA

Izolacja plazmidowego DNA była prowadzona z hodowli nocnej bakterii w pożywce LB z dodatkiem antybiotyku odpowiedniego dla genu oporności niesionego przez izolowany plazmid. Izolację prowadzono z wykorzystaniem komercyjnego zestawu Plasmid Mini AX (A&A Biotechnology), opartego na metodzie lizy alkalicznej.

6.2.5 Przygotowanie komórek kompetentnych

Komórki kompetentne przygotowano z wykorzystaniem komercyjnego zestawu Mix & Go *E. coli* Transformation Kit firmy Zymo Research zgodnie z protokołem producenta.

6.2.6 Transformacja komórek kompetentnych plazmidowym DNA

W przypadku transformacji plazmidami niosącymi gen oporności na ampicylinę, 100 µl zawiesiny komórek kompetentnych przygotowanych przy pomocy zestawu Mix & Go *E. coli* Transformation Kit (patrz wyżej) rozmrożono na lodzie, po czym dodano do niej 2 µl roztworu plazmidowego DNA. Komórki mieszano następnie przez ok. 5 s, wysiano na płytce z pożywką LA z dodatkiem ampicyliny a następnie hodowano w temperaturze 37°C przez noc.

Transformację plazmidami niosącymi gen oporności na kanamycynę przeprowadzono analogicznie, przy czym przed wysianiem na płytkę mieszaninę komórek kompetentnych i plazmidowego DNA dodano do 400 µl pożywki LB i hodowano z wytrząsaniem z prędkością 350 obr/min w temperaturze 37 °C przez 1h w termomikserze. Uzyskaną hodowlę następnie wirowano przez 2 min przy przyspieszeniu 5000 x g. Osadzone komórki zawieszono w 100 µl pożywki LB, a następnie wysiano na płytce z pożywką LA z dodatkiem kanamycyny. Wysiane komórki hodowano następnie w temperaturze 37 °C przez noc.

6.2.7 Nadprodukcja i oczyszczanie białek

6.2.7.1 IbpA_{E.coli}, domena ACD białka IbpA_{E.coli}, DnaK, DnaJ, GrpE, ClpB, lucyferaza ze znacznikiem polihistydynowym

Białka DnaK, DnaJ, GrpE, ClpB, IbpA z *E. coli*, domena ACD białka IbpA z *E. coli* jak i lucyferaza ze znacznikiem polihistydynowym zostały pozyskane z zasobów Pracowni Biochemii Białek. Były one oczyszczone w ramach wcześniejszych badań prowadzonych w Pracowni (Ratajczak, Zietkiewicz, i Liberek 2009; Obuchowski i in. 2019).

6.2.7.2 Białka IbpA

Nadprodukcja białek IbpA (IbpA_{E. amylovora}, IbpA_{E. coli} Q66H G109D, IbpA_{E. amylovora} H67Q D110G, AncA₀, AncA₁, AncA₀Alt, AncA₁Alt, AncA₀ Q66H, AncA₀ G109D, AncA₀ Q66H G109D, AncA₁ H66Q D109G AncA₀+7 oraz 7 mutantów białka AncA₀+7) była prowadzona w komórkach *E. coli* BL21 (DE3), transformowanych plazmidami pET3a zawierającymi geny kodujące dane białka.

Odmłodzona hodowla nocna *E. coli* była wytrząsana w pożywce LB z dodatkiem ampicyliny w 37 °C do momentu uzyskania przez nią wartości OD₆₀₀ w zakresie 0,4-0,8. Następnie nadprodukcja białka została zaindukowana przez dodanie do hodowli 0,5 mM IPTG.

Nadprodukcję prowadzono w tych samych warunkach przez 1,5h, po czym komórki oddzielono od pożywki przez wirowanie z przyspieszeniem 2957 x g przez 10 min.

Otrzymany osad komórek zawieszono następnie w buforze do lizy A (50 mM Tris pH 7,5, 50 mM NaCl, 5mM EDTA, 10% glicerol 5 mM β-merkaptetoetanol) wykorzystując homogenizator Pottera – Elvehjema, po czym lizowano przez sonikację w sonikatorze Qsonica (amplituda 13%, łączny czas sonikacji 2min 30s, podzielony na 15s impulsy oddzielone 45s przerwami).

Uzyskany w ten sposób lizat został następnie poddany wirowaniu z przyspieszeniem 75 600 x g przez 30 min w celu oddzielenia frakcji nierozpuszczalnej i rozpuszczalnej.

Frakcję nierozpuszczalną zawieszono w buforze A (40 mM Tris pH 7,5, 50 mM NaCl, 10% glicerol, 5 mM β – merkaptetoethanol, 6M mocznik) i inkubowano w nim z mieszaniem przez 15 min w celu ponownej solubilizacji, po czym ponownie wirowano z przyspieszeniem 75 600 x g przez 25 min w celu oddzielenia frakcji rozpuszczalnej w 6M moczniku od nierozpuszczalnego osadu.

Wykorzystując elektroforezę poliakrylamidową w warunkach denaturujących (SDS-PAGE) porównano następnie zawartość oczyszczanego białka w rozpuszczalnej frakcji lizatu i frakcji rozpuszczalnej w 6M moczniku.

Jeżeli oczyszczane białko znajdowało się głównie w rozpuszczalnej frakcji lizatu, zostało ono poddane wysalaniu przez dodanie $(\text{NH}_4)_2\text{SO}_4$ do stężenia 40 % i mieszanie przez co najmniej 1h. Wysolone białka zostały następnie oddzielone przez wirowanie z przyspieszeniem 75 600 x g przez 25 min i zawieszono w buforze A (patrz wyżej). W celu całkowitego usunięcia $(\text{NH}_4)_2\text{SO}_4$ i solubilizacji białek, roztwór dializowano następnie do czystego buforu A.

Jeżeli oczyszczone białko znajdowało się głównie we frakcji rozpuszczalnej w 6M moczniku, frakcja ta została bezpośrednio użyta w dalszych etapach oczyszczania.

Wyselekcjonowaną frakcję nałożono na kolumnę chromatograficzną ze złożem jonowymiennym Q - Sepharose, uprzednio zrównoważonym buforem A (patrz wyżej). Białka związane do złoża eluowano następnie gradientem stężenia NaCl w zakresie 50mM-500mM (ok. 8 objętości złoża), po czym zidentyfikowano frakcje zawierające pożądane białko przy pomocy elektroforezy poliakrylamidowej w warunkach denaturujących (SDS-PAGE). Wyselekcjonowane frakcje dializowano następnie do buforu B (40 mM Tris pH 8,5, 50 mM NaCl, 10% glicerol, 5 mM β - merkaptoetanol).

Jeżeli czystość uzyskanego roztworu białka nie była zadowalająca, został on dodatkowo nałożony na kolumnę chromatograficzną ze złożem jonowymiennym Q – Sepharose, zrównoważonym buforem B (patrz wyżej). Złoże płukano następnie buforem B, zbierając frakcję niezwiązaną do złoża (ang. *flow through*), zawierającą oczyszczane białko.

Po zakończeniu oczyszczania roztwór białka dializowano do buforu FL (50 mM Tris pH 7,5, 150 mM KCl, 10 % glicerol, 5 mM β - merkaptoetanol).

Czystość oczyszczonego białka zweryfikowano przy użyciu elektroforezy poliakrylamidowej w warunkach denaturujących, a jego stężenie wyznaczono przy użyciu metody Bradforda.

6.2.7.3 Białko IbpB z *E. coli*

Nadprodukcję współczesnego białka IbpB z *E. coli* prowadzono w komórkach *E. coli* BL21 (DE3), transformowanych plazmidem pET28b, zawierającym gen *ibpB*. Odmłodzoną hodowlę nocną w pożywce LB z dodatkiem kanamycyny wytrząsano w 37 °C do osiągnięcia wartości OD_{600} ok. 0,5, po czym indukowano nadprodukcję przez dodanie 1 mM IPTG. Nadprodukcję prowadzono przez 2 h, po czym hodowlę poddano wirowaniu z przyspieszeniem 2957 x g przez 8 min w celu oddzielenia komórek od pożywki.

Otrzymany osad komórek zawieszono w buforze do lizy B (40 mM Tris pH 7,4, 50 mM NaCl, 5mM EDTA, 10% glicerol 5 mM β -merkaptotoetanol) wykorzystując homogenizator Pottera – Elvehjema, po czym poddano lizie z wykorzystaniem prasy Frencha. Uzyskany lizat wirowano następnie z przyspieszeniem 75 600 x g przez 25 min, w celu oddzielenia frakcji rozpuszczalnej od nierozpuszczalnej.

Supernatant, stanowiący frakcję rozpuszczalną lizatu, poddano następnie wysalaniu przez dodanie $(\text{NH}_4)_2\text{SO}_4$ do stężenia 20% i mieszanie przez noc. Wysolone białka oddzielono następnie przez wirowanie z przyspieszeniem 75 600 x g przez 25 min, po czym zawieszono je w buforze C (40 mM Tris pH 7,5, 25 mM NaCl, 10% glicerol, 5 mM β – merkaptotoethanol, 6M mocznik). Uzyskaną zawiesinę dializowano następnie do czystego buforu C, w celu usunięcia $(\text{NH}_4)_2\text{SO}_4$ i ponownej solubilizacji wysolonych białek.

Próbkę nałożono następnie przez noc na kolumnę chromatograficzną zawierającą złożę jonowymienne Q - Sepharose zrównoważone buforem C. Kolumnę płukano buforem C, a następnie 20% buforu D (40 mM Tris pH 7,4, 500 mM NaCl, 10 % glicerol, 5 mM β – merkaptotoetanol, 6M mocznik) do momentu aż pomiar absorbancji frakcji opuszczającej kolumnę wykazał początek elucji białek.

Złożę inkubowano następnie przez 20 min przy zatrzymanym przepływie, po czym prowadzono elucję białek ze złoża w gradiencie NaCl (20% - 100% buforu D, 6 x objętość złoża). Frakcje zawierające białko lbpB zidentyfikowano przy użyciu elektroforezy poliakrylamidowej w warunkach denaturujących połączonej z barwieniem Coomassie Brilliant Blue.

Wyselekcjonowane frakcje dializowano do buforu E (40 mM Tris pH 8,5, 150 mM NaCl, 10 % glicerol, 5 mM β – merkaptotoetanol), a następnie do buforu FL (50 mM Tris pH 7,5, 150 mM KCl, 5% glicerol, 5 mM β -merkaptotoetanol).

Czystość oczyszczonego białka zweryfikowano przy użyciu elektroforezy poliakrylamidowej w warunkach denaturujących z barwieniem Coomassie Brilliant Blue, a jego stężenie wyznaczono przy użyciu densytometrii.

6.2.7.4 Domeny ACD białek AncA₀ i AncA₀ Q66H G109D

Nadprodukcja domen ACD była prowadzona w komórkach *E. coli* BL21 (DE3), transformowanych plazmidami pET28a zawierającymi geny kodujące domeny ACD z dołączonym znacznikiem heksahistydynowym.

Odmłodzona hodowla nocna *E. coli* była wytrząsana w pożywce LB z dodatkiem kanamycyny w 37 °C do momentu uzyskania przez nią wartości OD₆₀₀ w zakresie 0,4-0,8. Następnie nadprodukcja białka została zaindukowana przez dodanie do hodowli 0,5 mM IPTG.

Nadprodukcję prowadzono w tych samych warunkach przez 1,5h, po czym komórki oddzielono od pożywki przez wirowanie z przyspieszeniem 2957 x g przez 10 min.

Otrzymany osad komórek zawieszono następnie w buforze G (40 mM Tris pH 7,4, 500 mM NaCl, 10% glicerol, 5 mM β -merkaptotanol, 6M mocznik) wykorzystując homogenizator Pottera – Elvehjema, po czym lizowano przez sonikację w sonikatorze Qsonica (amplituda 13%, łączny czas sonikacji 2min 30s, podzielony na 15s impulsy oddzielone 45s przerwami).

Uzyskany lizat wirowano następnie z przyspieszeniem 75 600 x g przez 25 min, w celu oddzielenia frakcji nierozpuszczalnej w 6M moczniku. Supernatant, stanowiący frakcję rozpuszczalną lizatu, inkubowano przez noc ze złożem chromatograficznym Ni – NTA (Macherey-Nagel Protino NTA Agarose), zrównoważonym uprzednio buforem G.

Złoże było następnie umieszczone w kolumnie chromatograficznej i płukane buforem G i 5 % buforu H (40 mM Tris pH 7,4, 500 mM NaCl, 10% glicerol, 5 mM β -merkaptotanol, 360 mM imidazol, 6M mocznik). Po zakończeniu płukania białka zostały eluowane ze złoża gradientem imidazolu (5%-100% buforu H). Frakcje zawierające domeny ACD zostały zidentyfikowane z użycie elektroforezy poliakrylamidowej w warunkach denaturujących połączonej z barwieniem Coomasie Brilliant Blue.

Wyselekcjonowane frakcje zostały następnie dializowane do buforu I (50 mM Tris pH 8,5, 150 mM KCl, 10% glicerol, 5 mM β -merkaptotanol, 1 mM CaCl_2) i poddane trawieniu trombiną (0,5 U/ml próbki) w temperaturze pokojowej przez noc, w celu usunięcia znaczników heksahistydynowych.

Po zakończeniu trawienia trombina została inaktywowana przez dodanie 1mM PMSF, po czym próbka była dializowana do buforu J (50 mM Tris pH 8,5, 150 mM KCl, 10% glicerol, 5 mM β -merkaptotanol, 6M mocznik) i nałożona na kolumnę chromatograficzną ze złożem Ni-NTA zrównoważonym buforem J, w celu związania białek, których znaczniki heksahistydynowe nie uległy odcięciu.

Kolumnę płukano następnie buforem J, zbierając frakcję niezwiązaną do złoża (ang. *flow through*). Próbkę dializowano następnie do buforu FB' (50 mM Tris pH 8,5, 150 mM KCl, 5 mM β – merkaptotanol), a potem do buforu FB (50 mM Tris pH 7,5, 150 mM KCl, 5 mM β – merkaptotanol).

Czystość oczyszczonego białka zweryfikowano przy użyciu elektroforezy poliakrylamidowej w warunkach denaturujących z barwieniem Coomasie Brilliant Blue, a jego stężenie wyznaczono przy użyciu metody Bradforda.

6.2.7.5 His₆ – SUMO – peptyd C

Konstrukt His₆ – SUMO – peptyd C (region CTE białka AncA₀, o sekwencji PEAMKPPRIEIN z dołączonym znacznikiem SUMO i znacznikiem heksahistydynowym) był nadprodukowany w *E. coli* BL21 (DE3) transformowanych plazmidem pET28a zawierającym gen kodujący ten konstrukt.

Odmłodzona hodowla nocna była wytrząsana w 37 °C w pożywce LB z dodatkiem kanamycyny do momentu osiągnięcia wartości OD₆₀₀ ok. 0,5 – 1. Nadprodukcja została następnie indukowana przez dodanie 0,5 mM IPTG. Nadprodukcję prowadzono w tych samych warunkach przez 1,5h.

Po zakończeniu nadprodukcji hodowlę wirowano z przyspieszeniem 2957 x g przez 10 min. Uzyskany osad bakteryjny został następnie zawieszony w buforze K (40mM Tris pH 7,5, 100 mM NaCl, 10% glicerol, 5 mM β – merkaptoetanol, 10 mM imidazol) z wykorzystaniem homogenizatora Pottera – Elvehjema i poddany lizie przez sonikację w sonikatorze QSonica (amplituda 13%, łączny czas sonikacji 2min 30s, podzielony na 15s impulsy oddzielone 45s przerwami).

Uzyskany lizat następnie wirowano z przyspieszeniem 75 600 x g przez 25 min, w celu oddzielenia frakcji nierozpuszczalnej. Supernatant, stanowiący frakcję rozpuszczalną, inkubowano następnie ze złożem Ni – NTA (Macherey-Nagel Protino NTA Agarose), zrównoważonym buforem K (patrz wyżej) przez 1h. Złoże przeniesiono następnie do kolumny chromatograficznej i płukano buforem L (40 mM Tris pH 7,5, 100 mM NaCl, 10% glicerol, 5 mM β-merkaptoetanol, 40 mM imidazol), po czym eluowano białka związane do złoża przy użyciu buforu M (40 mM Tris pH 7,5, 100 mM NaCl, 10% glicerol, 5 mM β-merkaptoetanol, 250 mM imidazol).

Fracje zawierające oczyszczane białko zidentyfikowano przy użyciu elektroforezy poliakrylamidowej w warunkach denaturujących (SDS-PAGE), połączonej z barwieniem Coomassie Brilliant Blue. Wybrane frakcje dializowano następnie do buforu FL (50 mM Tris pH 7,5, 150 mM KCl, 10 % glicerol, 5 mM β – merkaptoetanol).

Czystość oczyszczonego białka zweryfikowano przy użyciu elektroforezy poliakrylamidowej w warunkach denaturujących z barwieniem Coomassie Brilliant Blue, a jego stężenie wyznaczono przy użyciu densytometrii.

6.2.8 Przygotowanie lizatu *E. coli*

Lizat *E. coli*, służący jako jeden z substratów w analizie oddziaływania sHsp i domen ACD ze zagregowanym substratem, został przygotowany przez Klaudię Kochanowicz. Do przygotowania lizatu wykorzystano szczep *E. coli* P_{IPTG} – *dua KJ ΔibpAB*, uzyskany dzięki uprzejmości dr. Axela Mogk, ZMBH, Universität Heidelberg.

Szczep ten posiada delecję operonu *ibpAB*, a ekspresja genów *dnaK* i *dnaJ* jest w nim kontrolowana przez promotor zależny od IPTG (Mogk i in. 2003).

Odmłodzoną hodowlę nocną wytrząsano w pożywce LB z dodatkiem 50 mM IPTG oraz kanamycyny i chloramfenikolu w temperaturze 25 °C. Po zakończeniu hodowli, komórki poddano lizie przez sonikację.

6.2.9 Oznaczanie stężenia białek

Stężenia białek IbpA oraz domen ACD jak i białek lizatu *E. coli* zostały oznaczone przy pomocy metody Bradforda. W celu uzyskania krzywej wzorcowej do 1 ml odczynnika Bradforda dodawano wzorcowy roztwór albuminy osocza wołowego (ang. *Bovine Serum Albumin*, BSA) w ilości 2 - 9 µg białka. Po 15 min inkubacji w temperaturze pokojowej zmierzono absorbancję próbek dla długości fali 595 nm przy użyciu spektrofotometru, po czym wyznaczono równanie prostej opisującej zależność między masą białka w próbce, a zmierzoną wartością absorbancji. Równolegle przeprowadzono analogiczny pomiar dla znanych objętości badanego roztworu białka i na podstawie wyznaczonego równania prostej wyznaczono jego stężenie.

Stężenia białka IbpB z *E. coli* jak i His₆-SUMO oraz konstruktu His₆ – SUMO – peptyd C wyznaczono przy użyciu densytometrii. Znane objętości badanych próbek jak i wzorcowy roztwór BSA zawierający rosnące masy białka zagotowano w 2x stężonym buforze Laemmli (62,5 mM Tris pH 6,8, 2,5 % SDS, 5 % β-merkaptoetanol, 10 % glicerol, błękit bromofenolowy) i poddano rozdziałowi elektroforetycznemu w warunkach denaturujących (SDS – PAGE), zgodnie ze standardową procedurą.

Żele poliakrylamidowe były następnie barwione barwnikiem Coomasie Brilliant Blue, a następnie odbarwione zgodnie ze standardową procedurą.

Po zakończeniu odbarwiania żele sfotografowano przy użyciu urządzenia ChemiDoc™ MP Imaging System. Wykorzystując program Image Lab 5.2 zmierzono następnie intensywność prążków pochodzących z badanego białka i porównano je z krzywą wzorową wyznaczoną na podstawie prążków pochodzących z wzorcowego roztworu BSA.

6.2.10 Reaktywacja lucyferazy

Pierwszym krokiem badania procesu dezagregacji i ponownego zwijania (reaktywacji) lucyferazy była jej denaturacja termiczna. 1,5 µM lucyferaza rozpuszczona w buforze F (150 mM KCl, 50 mM Tris pH 7,5, 20 mM MgCl₂, 2.5 mM DTT) była w tym celu inkubowana przez 10 min w 44 °C w termocyklerze. Denaturację lucyferazy prowadzono bez sHsp lub w obecności 10 µM sHsp (3 µM IbpA i 7 µM IbpB w przypadku systemu dwubiałkowego z *E. coli*).

Uzyskane w ten sposób agregaty lub kompleksy lucyferaza – sHsp były następnie inkubowane w temperaturze pokojowej w obecności systemu Hsp70 (1 μ M DnaK, 0,4 μ M DnaJ i 0,3 μ M GrpE), 2 μ M ClpB i systemu regeneracji ATP (5mM ATP, 0,1 mg/ml kinazy kreatyny i 18 mM fosforanu kreatyny). W jednym doświadczeniu zastosowano wyższe stężenie systemu Hsp70 (2 μ M DnaK, 0,8 μ M DnaJ i 0,6 μ M GrpE).

W poszczególnych punktach czasowych 2 μ l mieszaniny reakcyjnej dodawano do 10 μ l substratu lucyferazy Luciferase Assay System firmy Promega. Mierząc, przy pomocy luminometru GLOMAX™ 20/20, natężenie światła wydzielanego w trakcie reakcji katalizowanej przez lucyferazę, wyznaczono jej aktywność. Wyniki zostały przedstawione jako średni procent aktywności natywnej lucyferazy, wyznaczony z trzech powtórzeń \pm odchylenie standardowe.

6.2.11 Dynamiczne Rozpraszanie Światła (DLS)

Pomiary dynamicznego rozpraszania światła (ang. *Dynamic Light Scattering, DLS*) przeprowadzono z wykorzystaniem aparatu ZetaSizer Nano S firmy Malvern Instruments, dla 40 μ l próbki, przy zastosowaniu kąta rozpraszania 173° i długości fali 633 nm. Dla każdego pomiaru dokonano uśrednienia co najmniej 10 serii po 10 dziesięciosekundowych przebiegów i wyznaczono rozkład rozmiaru cząstek przez dopasowanie do 70 kategorii rozmiarów z zakresu 0,4 nm – 10 000 nm (Żwirowski i in. 2017).

6.2.11.1 Analiza odwracalnej deoligomeryzacji sHsp w warunkach szoku cieplnego

Aby zbadać zdolność sHsp do odwracalnej deoligomeryzacji w warunkach szoku cieplnego, wykorzystując pomiary DLS (patrz wyżej) określono rozmiar oligomerów utworzonych przez 10 μ M sHsp w buforze F (patrz wyżej). Pierwszy pomiar wykonano w temperaturze 25 °C, po czym próbkę dwukrotnie podgrzano do 44 °C (temperatury szoku cieplnego) i schłodzono do 25 °C. Po każdej zmianie temperatury ponownie zmierzono rozmiar tworzonych przez sHsp oligomerów przy użyciu DLS. Uzyskane wyniki przedstawiono jako rozkład średnic hydrodynamicznych oligomerów sHsp względem procentu zajmowanej przez nie objętości.

6.2.11.2 Analiza zdolności sHsp do sekwestracji agregujących substratów

W celu zmierzenia aktywności sekwestrazy badanych sHsp, denaturowano 1,5 μ M lucyferazy w buforze F (patrz wyżej) przez inkubację w 44°C przez 10 min w termocyklerze. Denaturacja była prowadzona w obecności rosnących stężeń sHsp lub w czystym buforze. Próbki zostały następnie schłodzone do 25 °C, a rozmiar utworzonych agregatów i/lub kompleksów lucyferaza – sHsp został zmierzony przy pomocy DLS (patrz wyżej). Wyniki przedstawiono jako średnią średnicę hydrodynamiczną uzyskanych cząstek \pm odchylenie standardowe.

6.2.11.3 Analiza wpływu substytucji na tworzenie oligomerów sHsp

W celu zbadania wpływu substytucji na pozycjach 66 i 109 białka AncA₀, wyznaczono przy użyciu pomiaru DLS rozmiar oligomerów tworzonych przez 10 μM białek AncA₀ i AncA₀ Q66H G109Dw buforze F (patrz wyżej) w 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 i 45°C. Wyniki zostały przedstawione jako rozkład średnicy hydrodynamicznej względem natężenia sygnału (dla temperatur 25, 35 i 45 °C) lub średniej średnicy hydrodynamicznej odpowiadającej maksimum dominującego wierzchołka rozkładu średnic hydrodynamicznych oligomerów względem procentu zajmowanej przez nie objętości.

6.2.12 Interferometria Biowarstwowa (BLI)

Do pomiarów z wykorzystaniem interferometrii biowarstwowej (ang. *Biolayer Interferometry* – BLI) użyto systemu Octet® K2 i ForteBio® BLItz. W pomiarach wykorzystano sensory Octet® NTA Biosensor, specyficznie wiążące znaczniki polihistydynowe.

Pomiary w systemie Octet® K2 prowadzono z wytrząsaniem z prędkością 1000 obr/min, w stałej temperaturze 25 lub 44°C, zależnie od doświadczenia. Pomiary w systemie ForteBio® BLItz prowadzono z wytrząsaniem z prędkością 2000 obr/min, w temperaturze pokojowej.

6.2.12.1 Analiza oddziaływania sHsp i domen ACD z agregatami białkowymi

Oddziaływanie sHsp i domen ACD z agregatami lucyferazy lub białek lizatu *E. coli* zostało zmierzone przy pomocy systemu Octet® K2.

Sensor BLI został pokryty kotwiczącą warstwą lucyferazy poprzez 10 min inkubację w 0,6 g/l lucyferazy ze znacznikiem heksahistydynowym w buforze F (patrz wyżej) z dodatkiem 4,5 M mocznika. Inkubację prowadzono w temperaturze pokojowej przy wytrząsaniu z prędkością 350 obr/min. Sensor został następnie przeniesiony do buforu FB (150 mM KCl, 50 mM Tris pH 7,5, 20 mM MgCl₂, 5 mM β-merkaptioetanol) i inkubowany kolejne 5 min w takich samych warunkach, w celu usunięcia mocznika i niezwiązanej lucyferazy.

Na tak przygotowanym sensorze utworzono następnie agregat białkowy przez 10 min inkubację w roztworze 0,5 g/l lucyferazy ze znacznikiem polihistydynowym lub 0,2 g/l białek frakcji rozpuszczalnej lizatu *E. coli*. Inkubację prowadzono w 44°C (w przypadku lucyferazy) lub w 55 °C (w przypadku białek lizatu) w termocyklerze. Sensor był następnie przeniesiony do czystego buforu FB (patrz wyżej) i inkubowany przez 5 min w temperaturze pokojowej z wytrząsaniem z prędkością 350 obr/min w celu usunięcia niezwiązanego białka.

Sensory z przyłączonym agregatem umieszczono w systemie Octet® K2 w czystym buforze FB i wykonano 60s pomiar linii bazowej. Sensory następnie

przeniesiono do 5 μ M roztworu badanych sHsp lub domen ACD w buforze FB i dokonano 1h pomiaru asocjacji sHsp do agregatu. Po zakończeniu tego etapu, sensory ponownie przeniesiono do czystego buforu FB w celu przeprowadzenia 1h pomiaru dysocjacji sHsp od agregatu. W przypadku wariantów alternatywnych sHsp pomiar asocjacji był prowadzony przez 10 min, a dysocjacji przez 60 min.

Pomiary prowadzono z wytrząsaniem z prędkością 1000 obr/min w temperaturze 44 °C (w przypadku białek sHsp) lub 25 °C (w przypadku domen ACD). Białka sHsp przed pomiarem były dodatkowo inkubowane przez 10 min w 44 °C w celu ich aktywacji.

6.2.12.2 Analiza oddziaływania domen ACD z regionem CTE sHsp

Oddziaływanie domen ACD z peptydem C – terminalnym również zostało zmierzone przy pomocy systemu Octet® K2. Sensor BLI Octet® NTA Biosensor umieszczono w buforze FL (150 mM KCl, 50 mM Tris pH 7,5, 5% glicerol, 5mM β -merkaptotoetanol) i dokonano 60s pomiaru linii bazowej. Sensor następnie umieszczono w 2,5 μ M roztworze peptydu C (PEMKPRRIEIN) z przyłączonym znacznikiem His₆ – SUMO w buforze FL i przeprowadzono asocjację peptydu do sensora przez 15 min. Niezwiązany peptyd został następnie usunięty przez inkubację sensora w buforze FL' (150 mM KCl, 50 mM Tris pH 7,5, 5mM β -merkaptotoetanol) przez 15 min. Po zakończeniu inkubacji sensor przeniesiono do roztworu domen ACD w buforze FL' (w różnych stężeniach) i przeprowadzono pomiar asocjacji domeny ACD do peptydu przez 48 min. Sensor następnie umieszczono w czystym buforze FL' i zmierzono dysocjację domeny ACD od peptydu przez 10 min.

Pomiary prowadzono z wytrząsaniem z prędkością 1000 obr/min w temperaturze 25 °C. Zostały one powtórzone trzy razy dla poszczególnych stężeń domen ACD. Wyniki przedstawiono jako zależność średniej grubości warstwy białek w momencie zakończenia etapu asocjacji od stężenia domeny ACD, po odjęciu pomiaru kontrolnego wykonanego pod nieobecność ACD.

6.2.12.3 Analiza inhibicji wiązania systemu Hsp70 do agregatów białkowych przez sHsp

Pomiar wiązania systemu Hsp70 do agregatów opłaszczonych sHsp wykonano przy użyciu urządzenia ForteBio® BLItz. Sensory z przyłączonymi agregatami lucyferazy przygotowano jak opisano powyżej (rozdział 6.2.12.1), wykorzystując bufor F zamiast buforu FB.

Pomiar linii bazowej prowadzono przez 60 s w buforze F. Po umieszczeniu sensorów w roztworze 5 μ M sHsp w buforze F, preinkubowanym w 44 °C przez 10 min, przeprowadzono pomiar asocjacji sHsp do agregatu przez kolejne 10 min. Sensor

przeniesiono następnie do czystego roztworu F lub do roztworu systemu Hsp70 i systemu regeneracji ATP w buforze F (0,7 μ M DnaK, 0,28 μ M DnaJ, 0,21 μ M GrpE, 5 mM ATP, 0,1 mg/ml kinazy kreatyny, 18 mM fosforanu kreatyny). Pomiar asocjacji systemu Hsp70 i/lub dysocjacji sHsp prowadzono następnie przez 60 min.

Wiązanie systemu Hsp70 do agregatu pod nieobecność sHsp zmierzono w analogiczny sposób, zastępując etap inkubacji w roztworze sHsp inkubacją w czystym buforze F.

6.2.13 Analiza selekcji naturalnej

Analiza selekcji naturalnej na wybranych gałęziach drzewa filogenetycznego została przeprowadzona w oparciu o analizę stosunku mutacji niesynonimicznych do synonimicznych (dN/dS, tzw. współczynnik ω). Wykorzystując zestawienie sekwencji aminokwasowych użyte do rekonstrukcji ancestralnej (rozdział 6.2.1) oraz odpowiadające im sekwencje nukleotydowe pozyskane z bazy danych NCBI, przygotowano zestawienie kodonowe przy pomocy platformy Pal2Nal. Z uzyskanego zestawienia kodonowego usunięto następnie pozycje o niskim wsparciu. Wykorzystując tak przygotowane zestawienie kodonowe, jak i drzewo filogenetyczne opisujące pokrewieństwo białek IbpA w *Enterobacterales* (rozdział 6.2.1), przeprowadzono analizę selekcji przy pomocy programu codeml.

W celu weryfikacji, czy wybrane gałęzie drzewa filogenetycznego znajdują się pod selekcją, przeprowadzono analizę opartą na modelu M0, zakładającym jednakowy stosunek ω dla wszystkich gałęzi (hipoteza zerowa) oraz na modelu „Two Ratio”, zakładającym, że wybrane gałęzie pierwszoplanowe (ang. *foreground*) mogą posiadać odmienny współczynnik ω niż reszta drzewa - tło (ang. *background*). Wyniki potwierdzono, wykonując dla wybranych gałęzi analizę opartą o model A1, zakładający, że pozycje na wybranej gałęzi mogą podlegać selekcji negatywnej ($\omega < 1$) lub dryfowi genetycznemu ($\omega = 1$) (hipoteza zerowa) oraz o model A, umożliwiającą dodatkowo obecność pozycji podlegających selekcji pozytywnej ($\omega > 1$).

Istotność statystyczną poszczególnych modeli wyznaczono w oparciu o test stosunku wiarygodności (ang. *Likelihood Ratio Test*) (Jeffares i in. 2015; Yang 1998; 2007; Yang i Nielsen 2002).

6.2.14 Analiza strukturalna i dynamika molekularna

Analiza *in silico* struktur dimerów domen ACD białek AncA₀ i AncA₀ Q66H G109D z zadokowanym regionem CTE białka AncA₀ (peptydem C) została przeprowadzona przez Marcina Pitka. Początkowa struktura kompleksu dimeru domen ACD białka AncA₀ z peptydem C została uzyskana przy użyciu platformy ColabFold (Mirdita i in. 2022), będącej implementacją programu AlphaFold2-Multimer, opartego o uczenie maszynowe (Evans i in. 2022). Substytucje Q66H i

G109D zostały następnie wprowadzone do tej struktury metodą mutagenyzy *in silico* z wykorzystaniem programu PyMol Mutagenesis Wizard (<http://www.pymol.org>).

Modele obu kompleksów zostały poddane następnie symulacji dynamiki molekularnej w stanie równowagi przez 500 ns z wykorzystaniem programu Gromacs 2019.2 (Van Der Spoel i in. 2005), stosując pole siłowe CHARMM36-jul2021 (Huang i in. 2017).

7. Wyniki

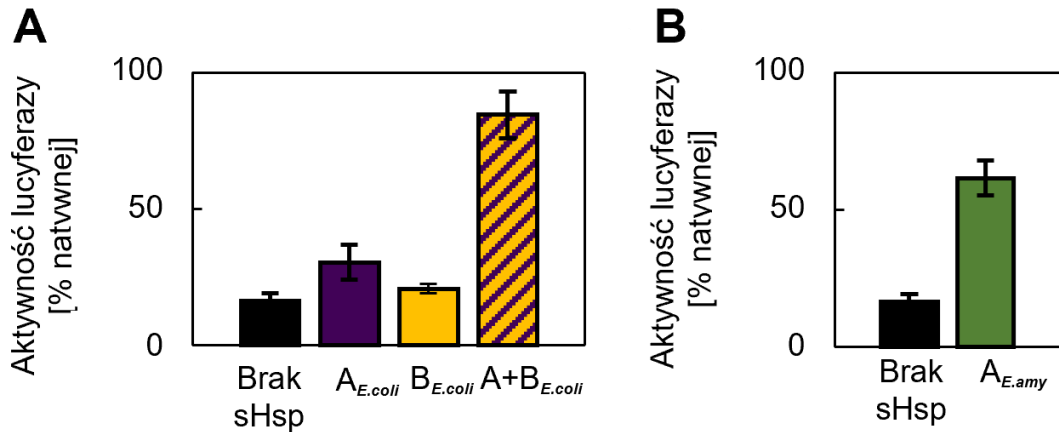
7.1 Porównanie zdolności współczesnych małych białek szoku cieplnego z *Erwiniaceae* i *Enterobacteriaceae* do wspomagania dezagregacji

Pierwszym krokiem było zmierzenie, w jakim stopniu współczesne sHsp z *Erwiniaceae* i *Enterobacteriaceae* wykorzystane w kolejnych doświadczeniach są zdolne do stymulacji dezagregacji i ponownego zwijania do struktury natywnej modelowego substratu przez system Hsp70. Jako modelowe białka należące do dwubiałkowego systemu z *Enterobacteriaceae* wykorzystałem białka IbpA i IbpB z *E. coli*, a jako model białka należącego do wtórnie pojedynczych sHsp z *Erwiniaceae* wykorzystałem białko IbpA z *E. amylovora*. Jako modelowy substrat wykorzystałem lucyferazę.

Lucyferaza jest enzymem często wykorzystywanym w badaniach procesu dezagregacji i ponownego zwijania białek. W konformacji natywnej katalizuje ona reakcję utleniania lucyferyny, czemu towarzyszy wytwarzanie światła. Umożliwia to szybki i precyzyjny pomiar aktywności tego enzymu przy użyciu luminometru. Pozwala to na obserwację dezagregacji i przywracania konformacji natywnej (reaktywacji) tego białka przez białka opiekuńcze.

Przeprowadziłem agregację 1,5 μM lucyferazy poprzez inkubację w 44 °C w obecności 10 μM sHsp (w przypadku systemu dwubiałkowego z *E. coli* użyłem 3 μM IbpA i 7 μM IbpB). Tak przygotowane agregaty inkubowałem następnie w temperaturze 25 °C w obecności bakteryjnego systemu Hsp70 (DnaK, DnaJ, GrpE) i białka Hsp100 (ClpB). Po godzinie inkubacji zmierzyłem aktywność lucyferazy w próbce przy pomocy luminometru.

W przypadku dwubiałkowego systemu sHsp z *E. coli* białko IbpA było w stanie jedynie w niewielkim stopniu samodzielnie stymulować reaktywację lucyferazy, podobnie białko IbpB - samodzielna stymulacja zachodziła w tym przypadku na poziomie porównywalnym z kontrolą pozbawioną sHsp. Kompletny system składający się z białek IbpA i IbpB był z kolei w stanie wydajnie stymulować reaktywację lucyferazy (Ryc. 8A). W przypadku wtórnie pojedynczego białka IbpA z *E. amylovora*, było ono w stanie wydajnie stymulować reaktywację lucyferazy pomimo braku białka IbpB (Ryc. 8B).



Ryc. 8 Stymulacja reaktywacji lucyferazy przez współczesne sHsp z *E. coli* i *E. amylovora*: 1,5 μ M lucyferaza była zagregowana przez inkubację z 10 μ M sHsp (3 μ M IbpA i 7 μ M IbpB w przypadku dwubiłkowego systemu z *E. coli*) w 44 °C, po czym reaktywowana przez inkubację z systemem Hsp70 i ClpB i systemem regeneracji ATP (1 μ M DnaK, 0,4 μ M DnaJ, 0,3 μ M GrpE, 2 μ M ClpB, 5 mM ATP, 0,1 mg/ml kinazy kreatyny, 18 mM fosforanu kreatyny) przez 1h. aktywność lucyferazy po reaktywacji została oznaczona przez luminometryczny pomiar katalizowanej przez nią reakcji utleniania lucyferyny. A) Dwubiłkowy system sHsp z *E. coli*. B) Wtórnie pojedyncze sHsp z *E. amylovora*.

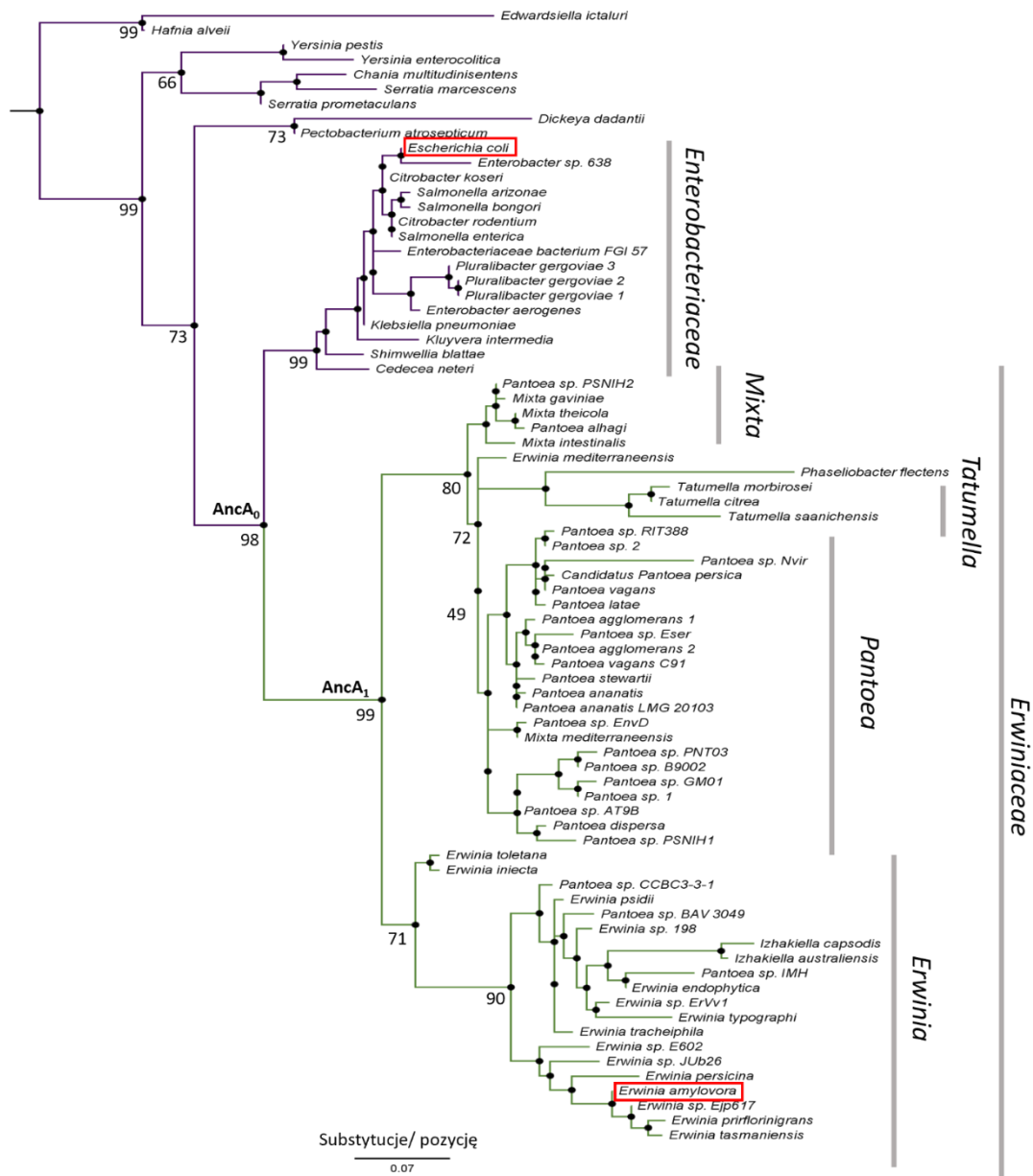
7.2 Analiza filogenetyczna białek IbpA w rzędzie *Enterobacterales*

Aby zbadać, w jaki sposób białko w IbpA w *Erwiniaceae* nabyło zdolność do wydajnej stymulacji dezagregacji bez udziału współpracującego białka IbpB, przeanalizowałem historię ewolucyjną białek IbpA w rzędzie *Enterobacterales*, w którego skład wchodzi m.in. rodziny *Erwiniaceae* i *Enterobacteriaceae*.

Wykorzystanie analizy historii ewolucyjnej, w odróżnieniu od prostego porównania właściwości i sekwencji współczesnych białek IbpA z tych dwóch rodzin, posiada dwie główne korzyści. Po pierwsze, pozwala na ograniczenie ilości analizowanych różnic przez eliminację zmian nabytych w toku ewolucji już po nabyciu nowej funkcji. Po drugie, taka analiza pozwala uniknąć sytuacji, w których zjawisko epistazy, czyli zależności efektu danej mutacji od szerszego kontekstu sekwencji aminokwasowej, uniemożliwia jednoznaczne wnioskowanie o genezie nowej funkcji na podstawie mutacji wprowadzonych we współczesnych białkach (Harms i Thornton 2010).

Pierwszym krokiem tej analizy była rekonstrukcja filogenezy molekularnej rodziny białek IbpA opisująca najbardziej wiarygodny scenariusz pokrewieństwa między współczesnymi homologami IbpA w *Enterobacterales*. Do inferencji filogenezy wykorzystałem program Iq-Tree, oparty na algorytmie maksymalnej wiarygodności (ang. *Maximum Likelihood*), znajdującym drzewo filogenetyczne z

największym prawdopodobieństwem prowadzące do danego zestawienia współczesnych sekwencji przy założeniu danego modelu ewolucji. Analiza została przeprowadzona na podstawie zestawienia 77 sekwencji aminokwasowych współczesnych białek IbpA z *Enterobacterales*, pochodzących z baz NCBI i UniProt. Przy użyciu programu Iq-Tree wyznaczyłem, w oparciu o test statystyczny BIC (ang. *Bayesian Information Criterion*), model ewolucji, który najlepiej opisuje zestawienie sekwencji białek IbpA (Kalyaanamoorthy i in. 2017; Nguyen i in. 2015). Model JTT+R3 najlepiej pasował do danych według przyjętego kryterium. Topologia uzyskanego drzewa filogenetycznego białek IbpA w *Enterobacterales* (Ryc. 9) charakteryzuje się wysokim wsparciem statystycznym (bootstrap) i pokrywa się z topologią wcześniej opublikowanych drzew filogenetycznych opisujących pokrewieństwo między poszczególnymi gatunkami należącymi do kładów *Erwiniaceae* i *Enterobacteriaceae* (Palmer i in. 2018; Obuchowski i in. 2019). Sekwencje IbpA z gatunków należących do rodziny *Erwiniaceae*, w których doszło do utraty partnerskiego białka IbpB, tworzą grupę monofiletyczną i pozostają w relacji siostrzanej do sekwencji IbpA z gatunków należących do rodziny *Enterobacteriaceae*.



Ryc. 9 Drzewo filogenetyczne białek IbpA w *Enterobacteriales*: Filogeneza została zrekonstruowana w oparciu o zestawienie 77 sekwencji ortologów IbpA z *Enterobacteriales* przy użyciu algorytmu maksymalnej wiarygodności (ang. Maximum Likelihood) przy założeniu modelu ewolucji JTT + R3. Na drzewie zaznaczono rodziny Enterobacteriaceae i Erwiniaceae, jak i rodzaje należące do rodziny Erwiniaceae; AncA₀ – węzeł reprezentujący ostatniego wspólnego przodka IbpA z Erwiniaceae i Enterobacteriaceae; AncA₁ – węzeł reprezentujący ostatniego wspólnego przodka IbpA z Erwiniaceae; wsparcie bootstrap zaznaczono dla kluczowych węzłów; współczesne ortologi IbpA z *E. coli* i *E. amylovora* zaznaczono czerwoną ramką; skala – substytucje/pozycję.

7.3 Rekonstrukcja ancestralnej sekwencji aminokwasowej białek AncA₀ i AncA₁

Wykorzystując zestawienie i drzewo filogenetyczne uzyskane w poprzednim etapie badań (rozdział 7.2), przeprowadziłem rekonstrukcję sekwencji aminokwasowych ancestralnych białek lbpA z *Enterobacteriales*. W analizie wykorzystałem program FastML, opierający się na algorytmie maksymalnej wiarygodności i empirycznym podejściu Bayesowskim (ang. *Empirical Bayes approach*). Wykorzystałem algorytm rekonstrukcji marginalnej (ang. *marginal reconstruction*), dla każdego z ancestralnych węzłów wyznaczając sekwencję aminokwasową, która z największym prawdopodobieństwem *a posteriori* mogła doprowadzić do powstania współczesnych sekwencji, przy założeniu określonego drzewa filogenetycznego i modelu ewolucji (Ashkenazy i in. 2012; Cohen i Pupko 2011; Cohen i in. 2008; Jones, Taylor, i Thornton 1992; Pupko i in. 2002; Simmons i Ochoterena 2000). Jako macierz substytucji przyjąłem macierz JTT, wyznaczoną na poprzednim etapie jako optymalnie opisującą zestawienie sekwencji użyte w analizie (rozdział 7.2).

Wynikiem mojej analizy był rozkład prawdopodobieństwa *a posteriori* dla każdej pozycji w zrekonstruowanej sekwencji białka. Najbardziej wiarygodna sekwencja dla danego węzła na każdej pozycji zawiera aminokwas o najwyższym prawdopodobieństwie *a posteriori*.

Aby zweryfikować hipotezę, że nabycie nowej funkcjonalności przez białko lbpA z *Erwiniaceae* nastąpiło równoległe z utratą białka lbpB, w dalszej analizie skupiłem się na dwóch ancestralnych sekwencjach – ostatnim wspólnym przodku białek lbpA z *Erwiniaceae* i *Enterobacteriaceae*, który wciąż występował w organizmie zawierającym białko lbpB (AncA₀), jak i ostatnim wspólnym przodku wtórnie pojedynczych białek lbpA z *Erwiniaceae*, występującym w organizmie, który utracił białko lbpB (AncA₁) (Ryc.9).

Na podstawie rekonstruowanych najbardziej wiarygodnych sekwencji można stwierdzić, że w kładzie *Erwiniaceae* między białkami AncA₀ i AncA₁ nastąpiło 10 zmian aminokwasowych, a między białkami AncA₁ i współczesnym lbpA z *E. amylovora* nastąpiły kolejne 24 zmiany. W kładzie *Enterobacteriaceae* z kolei między białkiem AncA₀ i współczesnym białkiem lbpA z *E. coli* zaszło 11 zmian.

Sprawia to, że identyfikacja różnic w sekwencji aminokwasowej odpowiedzialnych za różnice funkcjonalne jest potencjalnie łatwiejsza w przypadku porównywania białek AncA₀ i AncA₁, niż w przypadku porównywania współczesnych białek lbpA z *E. coli* i *E. amylovora*, których sekwencje aminokwasowe różnią się na 30 pozycjach.

Aby scharakteryzować właściwości biochemiczne białek AncA₀ i AncA₁, przeprowadziłem ich nadprodukcję w *E. coli*, wykorzystując plazmidy zawierające *de novo* syntetyzowane geny *ancA₀* i *ancA₁*, po czym oczyściłem je wykorzystując chromatografię jonowymienną.

7.4 Charakterystyka biochemiczna rekonstruowanych białek AncA₀ i AncA₁

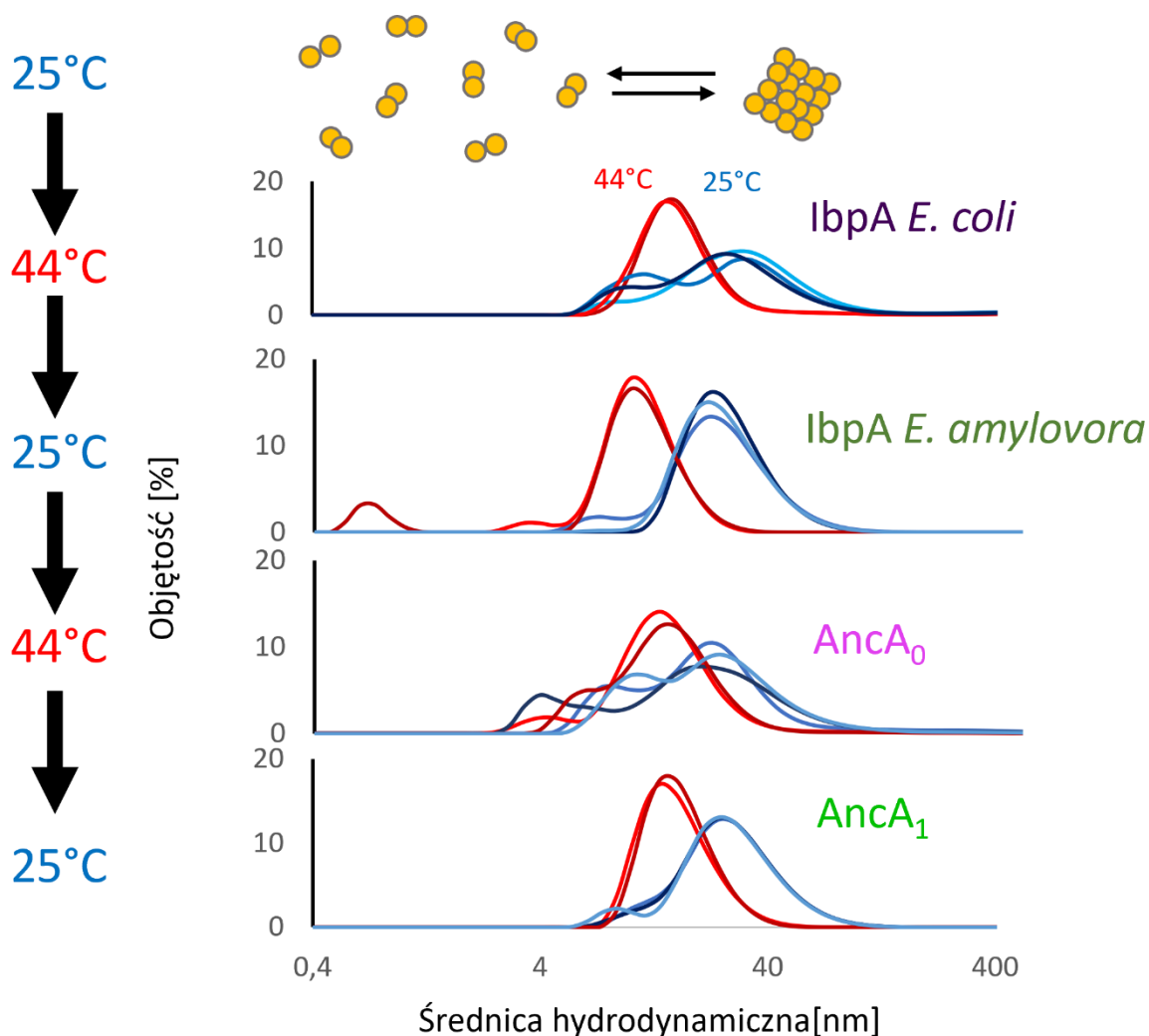
7.4.1 Analiza zdolności białek AncA₀ i AncA₁ do odwracalnej deoligomeryzacji w warunkach szoku cieplnego

Pierwszym krokiem analizy rekonstruowanych białek ancestralnych była weryfikacja, czy stanowią one funkcjonalne małe białka szoku cieplnego. Cechą charakterystyczną niemal wszystkich sHsp jest tworzenie dynamicznych oligomerów. W zależności od różnych czynników, takich jak modyfikacje posttranslacyjne lub zmiany warunków, sHsp mogą przechodzić w różne formy oligomeryczne (Haslbeck i Vierling 2015). W przypadku sHsp w *Enterobacterales* przybiera to formę odwracalnej deoligomeryzacji w warunkach szoku cieplnego, powiązanej z aktywacją tych białek (Ratajczak, Zietkiewicz, i Liberek 2009).

W celu weryfikacji, czy rekonstruowane białka AncA₀ i AncA₁ wykazują tę właściwość, wykorzystałem technikę Dynamicznego Rozpraszania Światła (DLS, z ang. *Dynamic Light Scattering*). Pozwala ona na pośredni pomiar średnicy hydrodynamicznej dyfundujących cząstek na podstawie prędkości ich dyfuzji.

Inkubowałem 10 μM roztwory białek AncA₀, AncA₁, jak i współczesnych białek IbpA z *E. coli* i *E. amylovora* naprzemiennie w temperaturze 25 °C i 44 °C, po każdej zmianie temperatury mierząc rozmiar tworzonych oligomerów przy pomocy DLS.

W przypadku wszystkich czterech analizowanych białek, podniesienie temperatury do 44 °C skutkowało wyraźnym zmniejszeniem obserwowanych oligomerów w porównaniu z pomiarem dokonany w 25 °C. Proces ten był całkowicie odwracalny po ponownym schłodzeniu, podgrzaniu i kolejnym schłodzeniu próbek (Ryc. 10).



Ryc. 10 Rekonstruowane ancestralne sHsp odwracalnie deoligomeryzują w warunkach szoku cieplnego podobnie do współczesnych sHsp: Rozkłady średnic hydrodynamicznych oligomerów tworzonych przez sHsp inkubowane naprzemiennie w 25°C (niebieskie wykresy) i 44°C (czerwone wykresy); średnice hydrodynamiczne wyznaczone za pomocą dynamicznego rozpraszania światła (DLS)

Wskazuje to, że oba rekonstruowane ancestralne białka wykazywały charakterystyczną dla bakteryjnych sHsp zdolność do odwracalnej deoligomeryzacji w warunkach szoku cieplnego, porównywalną do współczesnych białek IbpA z *E. coli* i *E. amylovora*.

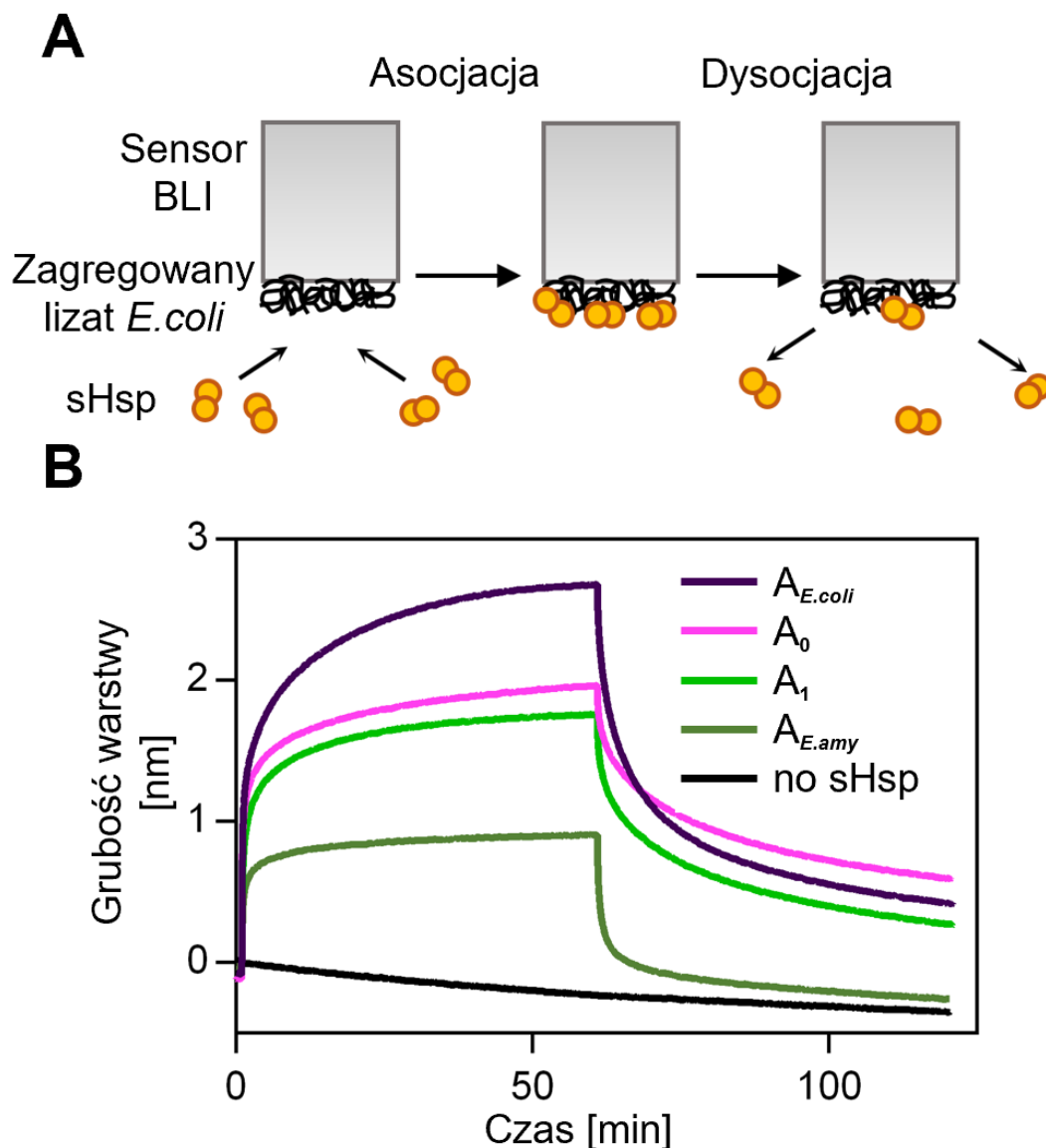
7.4.2 Analiza zdolności białek AncA₀ i AncA₁ do wiązania agregatów białkowych

Kolejną zbadaną przeze mnie właściwością rekonstruowanych białek ancestralnych było ich powinowactwo do zagregowanych substratów białkowych. Wykorzystałem w tym celu interferometrię biowarstwową (BLI, z ang. *Biolayer Interferometry*). Technika ta pozwala na precyzyjny pomiar zmian grubości warstwy biomolekuł związanych do sensora, dzięki pomiarowi interferencji wiązki światła

odbitej od granicy faz sensor - biomolekuły z wiązką światła odbitą od granicy faz biomolekuły - bufor.

Jako modelowy substrat wykorzystałem białka lizatu komórek *E. coli*, zagregowane przez inkubację w 44 °C z sensorem BLI pokrytym jonami niklu, z przyłączoną lucyferazą ze znacznikiem polihistydynowym (Obuchowski *et al.* 2019). Lizat wykorzystany w tym doświadczeniu został przygotowany przez Klaudię Kochanowicz. W celu ograniczenia wpływu białek opiekuńczych pochodzących z lizowanych komórek, użyty szczep *E. coli* zawierał delecję genów *ibpA* i *ibpB*, jak również obniżoną ekspresję genów *dnaJ* i *DnaK* (Mogk i in. 2003).

W celu pomiaru powinowactwa, inkubowałem analizowane sHsp z sensorem pokrytym zagregowanymi białkami przez godzinę, obserwując ich wiązanie do agregatu. Następnie przenieśliem sensor do czystego buforu na kolejną godzinę, obserwując dysocjację sHsp od agregatu. Pomiaru dokonałem w temperaturze szoku cieplnego (44 °C), w celu aktywacji sHsp (Ryc. 11).



Ryc. 11 Wiązanie rekonstruowanych i współczesnych białek IbpA do zagregowanego substratu: (A) Schemat doświadczenia; sensor BLI pokryty zagregowanymi białkami lizatu *E. coli* inkubowano w roztworze $5 \mu\text{M}$ sHsp przez 60 min, po czym płukano przez inkubację w czystym buforze przez 60 min. (B) zmiany grubości warstwy białek na sensorze BLI podczas asocjacji i dysocjacji rekonstruowanych i współczesnych sHsp.

Najszybsze wiązanie i największą grubość warstwy zaobserwowałem dla współczesnego białka IbpA z *E. coli*, podczas gdy najwolniejsze wiązanie i najmniejszą grubość utworzonej warstwy zaobserwowałem dla współczesnego białka IbpA z *E. amylovora*. Wskazuje to na dużo wyższe powinowactwo białka z *E. coli* do zagregowanego substratu w porównaniu do białka z *E. amylovora*, co pozostaje w zgodzie z wcześniejszymi badaniami (Obuchowski i in. 2019). W przypadku rekonstruowanych białek ancestralnych, oba wykazywały powinowactwo pośrednie

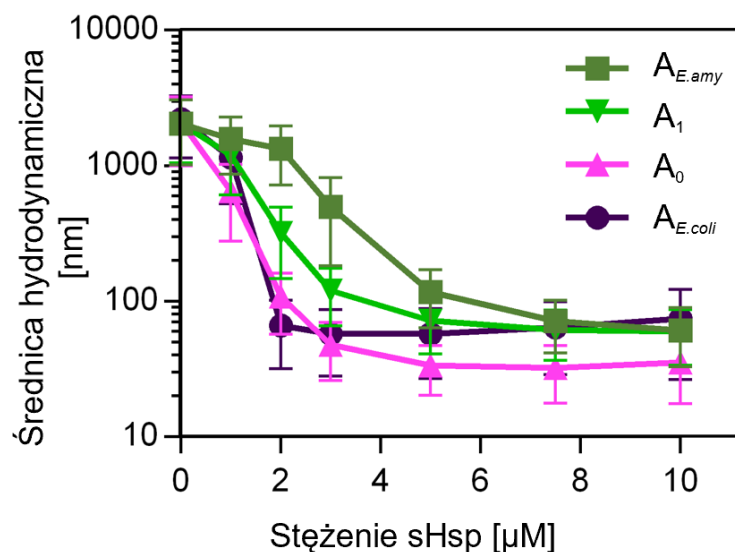
między współczesnymi białkami z *E. coli* i *E. amylovora*, przy czym białko AncA₀ wykazywało nieznacznie wyższe powinowactwo do zagregowanego lizatu w porównaniu z białkiem AncA₁ (Ryc. 11).

7.4.3 Analiza zdolności białek AncA₀ i AncA₁ do sekwestracji agregującego substratu

Podstawową aktywnością małych białek szoku cieplnego z rzędu *Enterobacterales* jest zdolność do sekwestracji agregujących substratów w kompleksach substrat – sHsp, zapobiegająca niekontrolowanej agregacji i potencjalnie ułatwiająca późniejszą dezagregację dzięki korzystniejszemu stosunkowi powierzchni takich kompleksów do ich objętości w porównaniu z agregatami utworzonymi bez udziału sHsp.

W celu zbadania aktywności sekwestrazy rekonstruowanych ancestralnych sHsp, ponownie wykorzystałem technikę DLS (rozdział 7.4.1). 1,5 μM modelowego substratu, lucyferazy, inkubowałem w warunkach szoku cieplnego (44°C) w celu doprowadzenia do jego agregacji. Inkubację prowadziłem w obecności rosnących stężeń badanych sHsp. Po zakończeniu inkubacji, zmierzyłem rozmiar utworzonych agregatów i/lub kompleksów sHsp – substrat.

Przy odpowiednio wysokim stężeniu sHsp, oba rekonstruowane ancestralne białka jak i współczesne białka IbpA z *E. coli* i *E. amylovora* były w stanie sekwestrować lucyferazę w kompleksach o dwa rzędy wielkości mniejszych niż agregaty utworzone przez nią pod nieobecność sHsp. Wskazuje to, że wszystkie analizowane białka posiadały aktywność sekwestrazy (Ryc. 12).



Ryc. 12 Sekwestracja agregującej lucyferazy przez rekonstruowane i współczesne sHsp: 1,5 μM lucyferazę inkubowano w 44°C przez 10 min w obecności rosnących stężeń sHsp, po czym zmierzono rozmiar utworzonych agregatów i/lub kompleksów sHsp – substrat przy pomocy DLS. Zmierzona średnica hydrodynamiczna przedstawiona jest jako wartość średnia \pm odchylenie standardowe.

Substrat był najwydajniej sekwestrowany przez współczesne białko lbpA z *E. coli* jak i rekonstruowane białko AncA₀. Białko AncA₁ wykazywało mniejszą zdolność do wydajnej sekwestracji agregującego substratu, podczas gdy współczesne białko lbpA z *E. amylovora* charakteryzowało się najmniejszą wydajnością sekwestracji substratu (Ryc. 12).

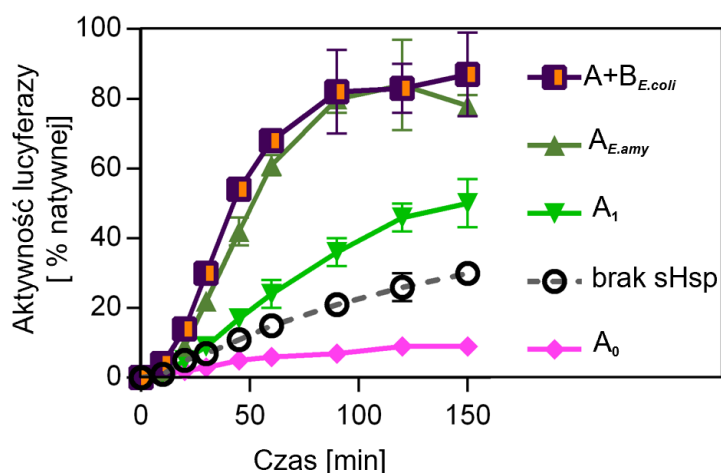
Wskazywałoby to, że w toku ewolucji wtórnie pojedynczego lbpA w *Erwiniaceae*, po utracie białka lbpB zdolność do wydajnej sekwestracji agregującego substratu przez białko lbpA zmalała.

7.4.4 Analiza zdolności białek AncA₀ i AncA₁ do wspomaganie dezagregacji i ponownego zwijania białek

Jak już zostało wspomniane, sekwestracja agregujących białek w kompleksach z sHsp może zwiększać wydajność ich późniejszej dezagregacji i ponownego zwijania do konformacji natywnej ze względu na mniejsze rozmiary takich kompleksów w porównaniu z agregatami jak i fakt, że sekwestrowane białka znajdują się w konformacji zbliżonej do natywnej. Z drugiej strony, sHsp opłaszczające sekwestrowane białka mogą utrudniać systemowi Hsp70 dostęp do substratów, tym samym utrudniając dezagregację, zwłaszcza przy limitujących stężeniach systemu Hsp70 (Żwirowski i in. 2017; Obuchowski i in. 2019).

W celu zbadania wpływu rekonstruowanych sHsp na proces dezagregacji i ponownego zwijania białek, przeprowadziłem test reaktywacji modelowego substratu – lucyferazy (rozdział 7.1), dokonując pomiarów aktywności lucyferazy w kolejnych punktach czasowych.

Współczesne, wtórnie pojedyncze białko lbpA z *E. amylovora*, jak i współczesny system dwóch białek lbpA i lbpB z *E. coli* charakteryzowały się porównywalną, wysoką efektywnością stymulacji reaktywacji lucyferazy. Ancestralne białko AncA₁ również stymulowało reaktywację lucyferazy w porównaniu z kontrolą nie zawierającą sHsp, jednak w mniejszym stopniu niż białka współczesne. Rekonstruowane białko AncA₀ z kolei nie tylko nie stymulowało reaktywacji, ale powodowało jej inhibicję w porównaniu z kontrolą (Ryc. 13). Najprawdopodobniej wiąże się to ze względnie silnym oddziaływaniem AncA₀ z substratem, znacząco utrudniającym dostęp do niego systemu Hsp70.



Ryc. 13 Zdolność rekonstruowanych i współczesnych sHsp do stymulacji reaktywacji lucyferazy:

Test reaktywacji lucyferazy przeprowadzono jak wyżej (ryc. 8); reaktywację prowadzono przez 150 min, oznaczając aktywność lucyferazy w poszczególnych punktach czasowych.

Powyższe wyniki wskazują, że w toku ewolucji wtórnie pojedynczego białka lbpA w *Erwiniaceae* zwiększyła się jego zdolność do stymulacji dezagregacji i ponownego zwijania białek, przy jednoczesnym spadku aktywności sekwestrazy. Zdolność do stymulacji tego procesu znacząco wzrosła (choć nie do poziomu białka współczesnego) pomiędzy węzłami AncA₀ i AncA₁, a więc równolegle do utraty białka lbpB w ostatnim wspólnym przodku *Erwiniaceae*.

7.5 Analiza zależności właściwości rekonstruowanych białek od niepewności statystycznej wynikającej z rekonstrukcji ancestralnej

Zastosowanie algorytmów opartych na maksymalnej wiarygodności pozwala na rekonstrukcję ancestralnych sekwencji aminokwasowych, które z największym prawdopodobieństwem doprowadziły do powstania obserwowanych współczesnych sekwencji, przy założeniu danego drzewa filogenetycznego i modelu ewolucji. Ze względu na probabilistyczny charakter analizy, rekonstruowana sekwencja obciążona jest jednak pewnym stopniem niepewności statystycznej, w szczególności dla pozycji charakteryzujących się niską konserwacją, wymagających wielu niezależnych substytucji do wyjaśnienia aminokwasów występujących na danej pozycji w sekwencjach współczesnych białek (Eick i in. 2016).

W związku z powyższym, wnioskowanie o historii ewolucyjnej na podstawie właściwości rekonstruowanych białek wymaga weryfikacji, w jakim stopniu właściwości te zależą od pozycji, których rekonstrukcja obciążona jest wysoką niepewnością (Eick i in. 2016).

W celu weryfikacji, w jakim stopniu właściwości białek AncA₀ i AncA₁ zależą od niepewnych pozycji, wykorzystałem metodę opartą o analizie tzw. najgorszego możliwego scenariusza (*ang. worst case scenario*) – analizę właściwości białka, w którym dla każdej niepewnej pozycji (zdefiniowanej jako pozycja, w której więcej niż jeden aminokwas mógł występować z prawdopodobieństwem *a posteriori* większym lub równym niż 0,2) najbardziej wiarygodny aminokwas zastąpiony jest najmniej wiarygodnym aminokwasem o prawdopodobieństwie *a posteriori* większym lub równym 0,2 (Eick i in. 2016).

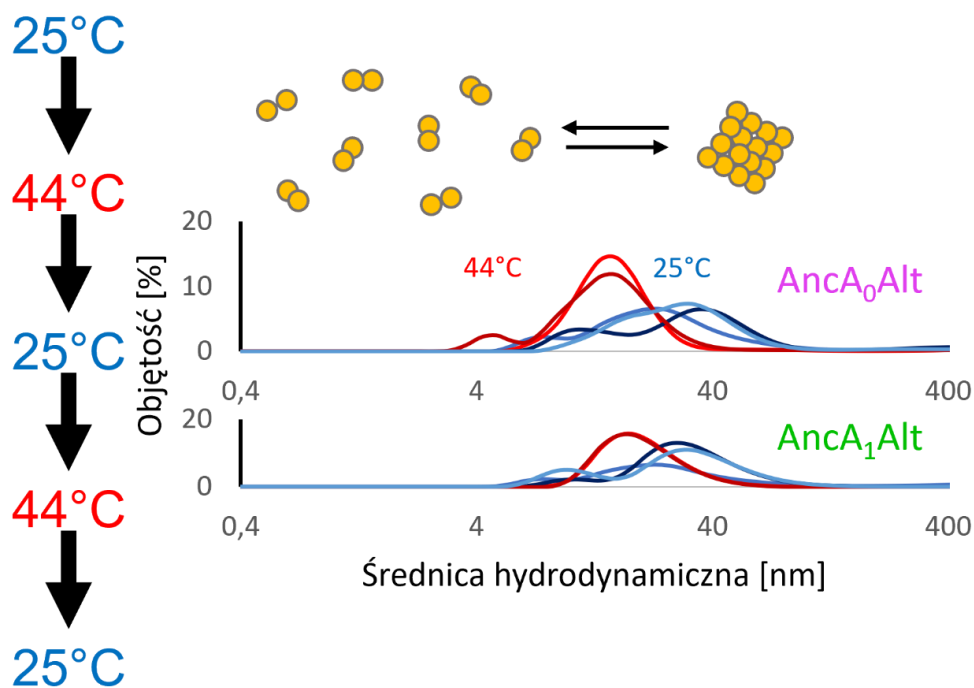
Białko AncA₀ zawierało cztery tak zdefiniowane niepewne pozycje (28, 66, 124, 137), a białko AncA₁ – dwie (124, 138). Na tej podstawie, zgodnie z przedstawioną powyżej zasadą, zaprojektowałem dwa „alternatywne” białka: AncA₀N28S Q66H V124I N137K, dalej określane jako AncA₀Alt i AncA₁V124I K138S, dalej określane jako AncA₁Alt.

W celu analizy w jakim stopniu właściwości uzyskanych w ten sposób wariantów przypominają właściwości białek o najbardziej wiarygodnych rekonstruowanych sekwencjach aminokwasowych, przeprowadziłem nadprodukcję i izolację białek AncA₀Alt i AncA₁Alt, po czym scharakteryzowałem ich zdolność do odwracalnej deoligomeryzacji w warunkach szoku cieplnego, wiązania do agregatów białkowych, sekwestracji agregujących substratów i wspomagania reaktywacji zagregowanej lucyferazy przez system Hsp70 i Hsp100.

Jako pierwszą przeprowadziłem analizę zdolności białek AncA₀Alt i AncA₁Alt do odwracalnej deoligomeryzacji w warunkach szoku cieplnego. Podobnie jak w przypadku analizy najbardziej wiarygodnych wariantów, wykorzystałem technikę DLS

w celu określenia średnicy hydrodynamicznej oligomerów tworzonych przez 10 μM roztwór małych białek inkubowanych naprzemiennie w 25 i 44 $^{\circ}\text{C}$ (rozdział 7.4.1).

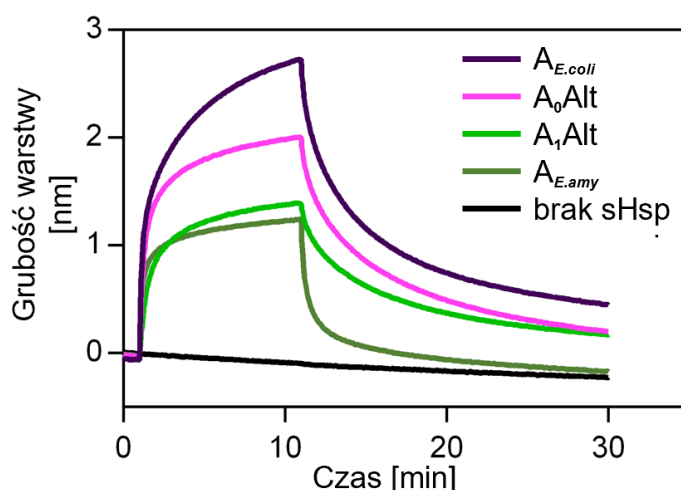
W przypadku białek AncA₀Alt i AncA₁Alt zaobserwowałem zmniejszenie rozmiaru tworzonych oligomerów po inkubacji w 44 $^{\circ}\text{C}$ i ponowne zwiększenie ich rozmiarów po inkubacji w 25 $^{\circ}\text{C}$, powtarzające się przy kolejnym podgrzaniu i schłodzeniu próbki, podobnie jak w przypadku najbardziej wiarygodnych wariantów AncA₀ i AncA₁ (rozdział 7.4.1) (Ryc. 14). Wskazuje to, że niepewność statystyczna rekonstrukcji ancestralnej nie ma znaczącego wpływu na zdolność białek AncA₀ i AncA₁ do odwracalnej deoligomeryzacji w warunkach szoku cieplnego.



Ryc. 14 Rekonstruowane alternatywne ancestralne sHsp odwracalnie deoligomeryzują w warunkach szoku cieplnego: Rozkłady średnic hydrodynamicznych oligomerów tworzonych przez sHsp inkubowane naprzemiennie w 25 $^{\circ}\text{C}$ (niebieskie wykresy) i 44 $^{\circ}\text{C}$ (czerwone wykresy). Średnice hydrodynamiczne wyznaczone za pomocą dynamicznego rozpraszania światła (DLS)

Następnie zbadałem zdolność wariantów AncA₀Alt i AncA₁Alt do wiązania zagregowanego substratu. Podobnie jak w przypadku wariantów o najwyższej wiarygodności, wykorzystałem technikę BLI. Inkubowałem analizowane białka z sensorem pokrytym zagregowanymi białkami lizatu *E. coli* (rozdział 7.4.2) przez 10 min, obserwując wzrost grubości warstwy białek na sensorze na skutek wiązania sHsp do agregatu. Następnie umieściłem sensor w czystym buforze przez 60 min, obserwując spadek grubości tej warstwy na skutek dysocjacji sHsp od agregatu.

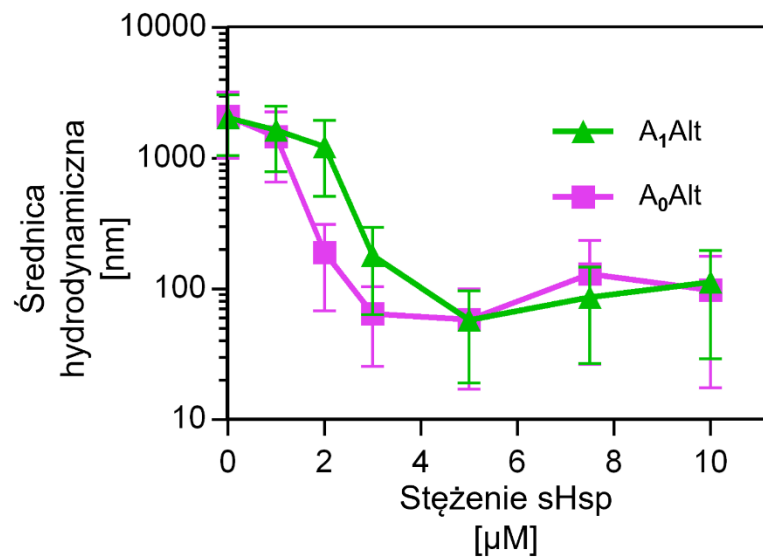
Podobnie jak w przypadku najbardziej wiarygodnych wariantów, białka AncA₀Alt i AncA₁Alt wykazywały powinowactwo do agregatów białkowych wyższe od współczesnego IbpA z *E. amylovora* i niższe od współczesnego IbpA z *E. coli*. Białko AncA₀Alt wykazywało wyższe powinowactwo niż białko AncA₁Alt (Ryc. 15). Wynika z tego, że niepewność statystyczna rekonstrukcji ancestralnej ma niewielki wpływ na zdolność białek AncA₀ i AncA₁ do wiązania zagregowanych substratów.



Ryc. 15 Rekonstruowane alternatywne ancestralne sHsp wiążą agregowany substrat: Zmiany grubości warstwy białek na sensorze BLI podczas asocjacji i dysocjacji alternatywnych rekonstruowanych sHsp jak i współczesnych sHsp. Eksperyment przeprowadzono jak uprzednio (ryc. 11), prowadząc asocjację przez 15 min, a dysocjację przez 60 min.

W celu analizy zdolności wariantów AncA₀Alt i AncA₁Alt do sekwestracji agregujących substratów w kompleksach sHsp – substrat, ponownie wykorzystałem technikę DLS. Przeprowadziłem agregację lucyferazy w obecności rosnących stężeń badanych sHsp poprzez inkubację w 44 °C, po czym zmierzyłem rozmiar utworzonych kompleksów i/lub agregatów przy pomocy DLS.

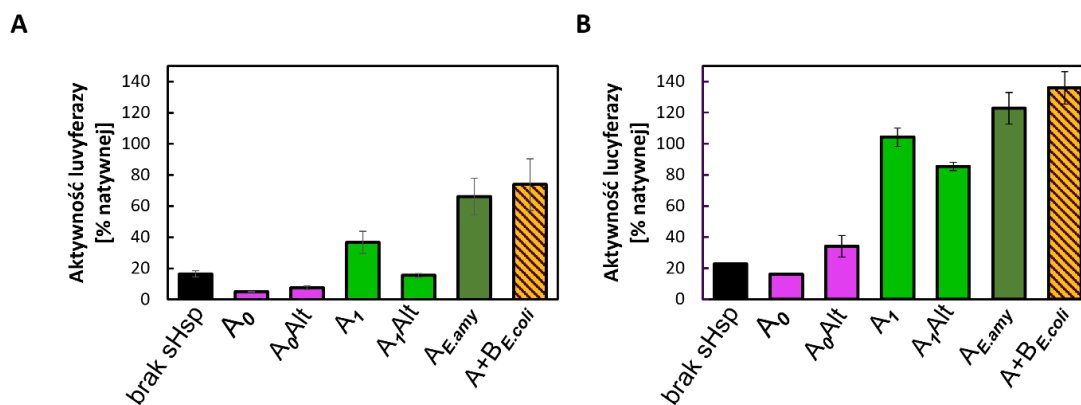
Podobnie jak warianty o najwyższej wiarygodności (rozdział 7.4.3), białka AncA₀Alt i AncA₁Alt były w stanie sekwestrować agregującą lucyferazę, przy czym AncA₀Alt charakteryzowało się większą wydajnością sekwestracji agregującego substratu niż AncA₁Alt (Ryc. 16). Wynika z tego, że niepewność związana z rekonstrukcją ancestralną ma niewielki wpływ na zdolność białek AncA₀ i AncA₁ do sekwestracji agregującego substratu.



Ryc. 16 Zdolność rekonstruowanych alternatywnych wariantów białek ancestralnych do sekwestracji agregującej lucyferazy: Aktywność sekwestrazy rekonstruowanych alternatywnych białek ancestralnych została zmierzona jak uprzednio (ryc. 12).

W celu analizy zdolności wariantów AncA₀Alt i AncA₁Alt do stymulacji procesu dezagregacji i ponownego zwijania białek, wykorzystałem test reaktywacji lucyferazy. Lucyferazę, zagregowaną w 44 °C w obecności 10 μM sHsp, inkubowałem przez godzinę w obecności bakteryjnego systemu Hsp70 i białka Hsp100 w temperaturze 25 °C. Po upływie godziny zmierzyłem aktywność lucyferazy w próbce przy pomocy luminometru (rozdział 7.4.4).

W przypadku, gdy reaktywacja prowadzona była przy takim samym stężeniu systemu Hsp70, jak użyty do pomiaru właściwości wariantów o największej wiarygodności (1 μM DnaK, 0,4 μM DnaJ, 0,3 μM GrpE), białko AncA₀Alt, podobnie jak białko AncA₀, inhibowało proces dezagregacji. Białko AncA₁Alt z kolei, w przeciwieństwie do wariantu AncA₁, nie było w stanie wydajnie stymulować procesu dezagregacji. W przypadku zastosowania dwukrotnie wyższego stężenia systemu Hsp70 (2 μM DnaK, 0,8 μM DnaJ, 0,6 μM GrpE), białko AncA₁Alt zaczęło wykazywać zdolność do stymulacji dezagregacji jedynie nieznacznie niższą niż AncA₁, a białko AncA₀Alt stymulowało dezagregację lucyferazy jedynie w niewielkim stopniu (Ryc. 17 A, B). Zgadza się to z wynikami wcześniejszych badań, pokazujących, że zdolność sHsp do stymulacji lub inhibicji dezagregacji zależy w znacznej mierze od stężenia systemu Hsp70 (Obuchowski i in. 2019; Źwirowski i in. 2017).



Ryc. 17 Zdolność alternatywnych i najbardziej wiarygodnych wariantów rekonstruowanych sHsp do stymulacji dezagregacji lucyferazy przez system Hsp70 i Hsp100: Test reaktywacji lucyferazy był przeprowadzony jak uprzednio (ryc. 8). A) Reaktywacja prowadzona przy standardowym stężeniu systemu Hsp70 (1 μ M DnaK, 0,4 μ M DnaJ, 0,3 μ M GrpE). B) Reaktywacja prowadzona przy wysokim stężeniu systemu Hsp70 (2 μ M DnaK, 0,8 μ M DnaJ, 0,6 μ M GrpE).

Wynika z tego, że, jakkolwiek niepewne pozycje w rekonstruowanych sekwencjach aminokwasowych mają wpływ na zdolność do stymulacji procesu dezagregacji i ponownego zwijania białek w niższych stężeniach systemu Hsp70 (w szczególności w przypadku białka AncA₁), obserwowany znaczący wzrost zdolności białka AncA₁ do stymulacji tego procesu w porównaniu z białkiem AncA₀ nie zależy w znaczący sposób od niepewności związanej z rekonstrukcją ancestralną.

Podsumowując, niepewność statystyczna związana z rekonstrukcją ancestralną nie ma wpływu na obserwowany wzrost zdolności do stymulacji dezagregacji i ponownego zwijania białek przy jednoczesnym zmniejszeniu aktywności sekwestrazy pomiędzy węzłami AncA₀ i AncA₁.

7.6 Analiza selekcji naturalnej pomiędzy węzłami AncA₀ i AncA₁

Aby określić mechanizm ewolucyjny, który najprawdopodobniej doprowadził do wykształcenia obserwowanej nowej aktywności białka AncA₁, zbadałem selekcję naturalną działającą na gałęzi drzewa filogenetycznego między węzłami AncA₀ i AncA₁. Zastosowana analiza opierała się na wyznaczeniu stosunku substytucji synonimicznych i niesynonimicznych (tzw. współczynnik ω).

Opierając się na wyznaczonym wcześniej drzewie filogenetycznym (rozdział 7.2), zestawieniu sekwencji aminokwasowych ortologów lbpA z *Enterobacterales* (rozdział 7.2) oraz odpowiadającym im sekwencjom nukleotydowym pozyskanych z bazy danych NCBI, wykazałem, że gałąź między węzłami AncA₀ i AncA₁ charakteryzuje wartość współczynnika ω większa od 1, co wskazuje, że nabycie nowych funkcji przez białko AncA₁ nastąpiło najprawdopodobniej na skutek działania selekcji pozytywnej.

Wynik ten został uzyskany zarówno przy wykorzystaniu modelu zakładającego jednakową selekcję dla wszystkich pozycji w sekwencji aminokwasowej na danej gałęzi (ang. *branch model*) jak i modelu zakładającego różną selekcję dla poszczególnych pozycji w sekwencji aminokwasowej na danej gałęzi (ang. *branch-site model*) (Jeffares i in. 2015; Yang 1998; 2007; Yang i Nielsent 2002).

model	LP	lnL	ω	LRT	
M0	154	-11631,91	$\omega=0,04$	-	
Two – Ratio, gałąź A ₀ -A ₁ jako pierwszoplanowa	155	-11618,80	$\omega_0=0,04$ $\omega_1=999,00$	LRT = 26,22	P<0,001
Two – Ratio, kład <i>Erwiniaceae</i> jako pierwszoplanowy	155	-11630,36	$\omega_0=0,04$ $\omega_1=0,04$	LRT = 3,11	NS

Tabela 1. Parametry statystyczne modeli selekcji typu *branch model*: W analizie model M0 stanowił hipotezę zerową; LP – liczba parametrów, lnL – logarytm naturalny wiarygodności (ang. Likelihood), ω – stosunek substytucji niesynonimicznych do synonimicznych; ω_0 – współczynnik ω dla gałęzi tła, ω_1 – współczynnik ω dla gałęzi pierwszoplanowych, LRT – test stosunku wiarygodności (ang. Likelihood Ratio Test), NS – nieistotny statystycznie

model	LP	lnL	Klasa pozycji	proporcja	ω_0	ω_1	LRT	
gałąź A ₀ -A ₁ jako pierwszoplanowa	A1	156	-11576,80	0	0,35	0,03	-	
				1	0,01	1,00		
				2a	0,63	0,03		
				2b	0,01	1,00		
	A	157	-11565,20	0	0,88	0,03	LRT = 23,20	P<0,001
				1	0,01	1,00		
				2a	0,10	0,03		
				2b	0,00	1,00		
kład <i>Erwiniaceae</i> jako pierwszoplanowy	A1	156	-11523,90	0	0,90	0,03	-	
				1	0,03	1,00		
				2a	0,07	0,03		
				2b	0,00	1,00		
	A	157	-11570,20	0	0,98	0,03	LRT = - 92,51	NS
				1	0,01	1,00		
				2a	0,01	0,03		
				2b	0,00	1,00		

Tabela 2. Parametry statystyczne modeli selekcji typu *branch – site model*: W analizie model A1 stanowił hipotezę zerową; LP – liczba parametrów, lnL – logarytm naturalny wiarygodności (ang. Likelihood), ω – stosunek substytucji niesynonimicznych do synonimicznych; ω_0 – współczynnik ω dla gałęzi tła, ω_1 – współczynnik ω dla gałęzi pierwszoplanowych, LRT – test stosunku wiarygodności (ang. Likelihood Ratio Test), NS – nieistotny statystycznie. Klasy pozycji: 0 – pozycje podlegające negatywnej selekcji zarówno w gałęziach tła, jak i pierwszoplanowych, 1 – pozycje podlegające dryfowi genetycznemu zarówno w gałęziach tła, jak i pierwszoplanowych, 2a – pozycje, które w gałęziach pierwszoplanowych podlegają selekcji pozytywnej, a w gałęziach tła – negatywnej, 2b – pozycje, które w gałęziach pierwszoplanowych podlegają selekcji pozytywnej, a w gałęziach tła – dryfowi genetycznemu

7.7 Identyfikacja kluczowych mutacji odpowiadających za różnice w aktywności białek AncA₀ i AncA₁

W celu identyfikacji substytucji aminokwasowych odpowiadających za obserwowane zmiany funkcjonalne między węzłami AncA₀ i AncA₁, porównałem sekwencje aminokwasowe obu rekonstruowanych białek. Różniły się one na dziesięciu pozycjach, z których osiem znajdowało się w domenie ACD, a dwie w regionie CTE (Ryc. 18).

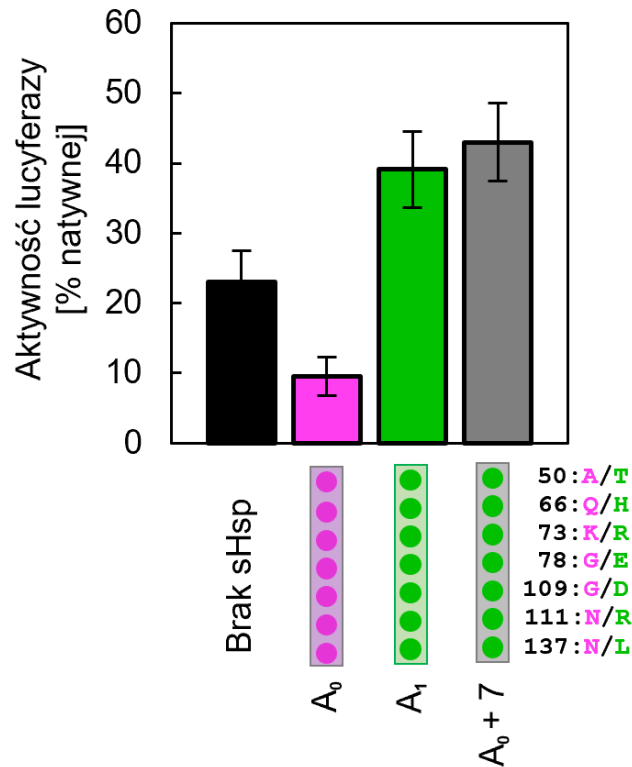
	region NTE	domena ACD	
AncA ₀	MRNFDLSPLYRSAIGFDRFLFNLLSNQNSGGYPYNNVELVDENHYRI	AI AVAGFAESEL DIT AQDNLLIVKGAHAGEQ	80
AncA ₁	MRNFDLSPLYRSAIGFDRFLFNLLSNQNSGGYPYNNVELVDENHYRI	TI AVAGFAQSEL DIT SHDNLLIVRGAAHEEQ	80
AncA ₀	PERTYLYQGI AERNFERKFQ LAEHIHVR	GANLENGLLYIDL ERVVPEAMKPRRIEIN-	137
AncA ₁	PERTYLYQGI AERNFERKFQ LAEHIHVR	DARLENGLLYIDL ERVVPEAMKPRRIEILK	138
	domena ACD	region CTE	

Ryc. 18 Zestawienie sekwencji aminokwasowych rekonstruowanych białek AncA₀ i AncA₁: Różnice wykorzystane w dalszych analizach zaznaczono na czerwono, różnice pominięte w dalszych analizach ze względu na niską konserwację we współczesnych ortologach *lbpA* zaznaczono przez pogrubienie.

Jako że największe różnice w aktywności białek AncA₀ i AncA₁ zaobserwowałem w przypadku pomiaru ich zdolności do stymulacji procesu dezagregacji i ponownego zwijania zagregowanych białek, do analizy wpływu poszczególnych substytucji na właściwości białek ancestralnych wykorzystałem test reaktywacji lucyferazy (rozdział 7.4.4).

Ze względu na niski stopień konserwacji we współczesnych białkach *lbpA* z rodziny *Erwiniaceae*, w dalszych analizach pominąłem trzy z dziesięciu potencjalnych pozycji (58, 65 i 138) (Ryc. 18). Aby potwierdzić, że substytucje na pozostałych siedmiu pozycjach są wystarczające, aby białko AncA₀ nabyło właściwości białka AncA₁, zaprojektowałem białko AncA₀ + 7, wprowadzając w białku AncA₀ na każdej z tych siedmiu pozycji substytucję na aminokwas występujący w białku AncA₁.

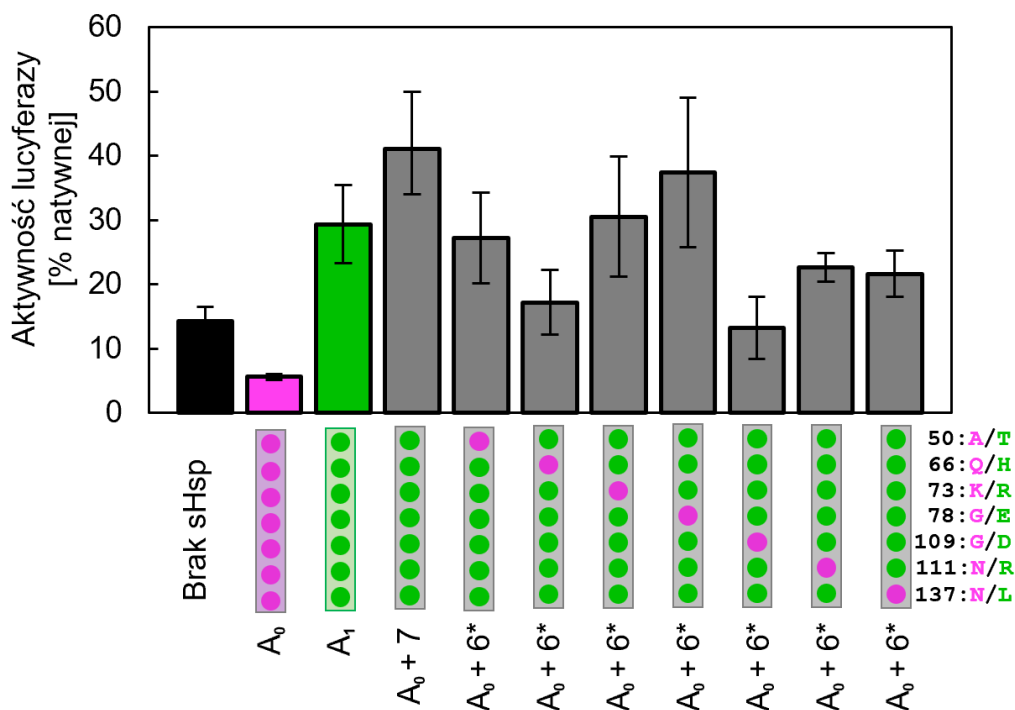
Białko AncA₀ + 7 było w stanie wydajnie stymulować proces reaktywacji lucyferazy, na poziomie porównywalnym z białkiem AncA₁ (Ryc. 19). Wskazuje to, że wyselekcjonowane siedem substytucji rzeczywiście wystarcza, aby białko AncA₀ nabyło właściwości białka AncA₁.



Ryc. 19 Substytucje na 7 wyselekcjonowanych pozycjach wystarczają, aby AncA₀ nabyło zdolność do stymulacji reaktywacji lucyferazy na poziomie AncA₁: Test reaktywacji lucyferazy został przeprowadzony jak uprzednio (ryc. 8); A₀+7 – białko AncA₀ z wprowadzonymi na pozycjach 50, 66, 73, 78, 109, 111 i 137 substytucjami na aminokwas występujący w białku AncA₁; substytucje zostały symbolicznie przedstawione na schemacie – pozycje oznaczone kolorem magenta oznaczają aminokwas występujące w bardziej ancestralnym białku AncA₀, a oznaczone kolorem zielonym – występujące w bardziej współczesnym białku AncA₁.

Aby określić, które z tych siedmiu substytucji są niezbędne dla uzyskania nowej funkcjonalności przez białko AncA₁, przygotowałem siedem kolejnych mutantów białka AncA₀. Były one zaprojektowane analogicznie do AncA₀ + 7, przy czym jedynie sześć z siedmiu analizowanych pozycji zawierało aminokwas charakterystyczny dla białka AncA₁, a jedna z nich, inna w przypadku każdego wariantu, zawierała aminokwas charakterystyczny dla białka AncA₀.

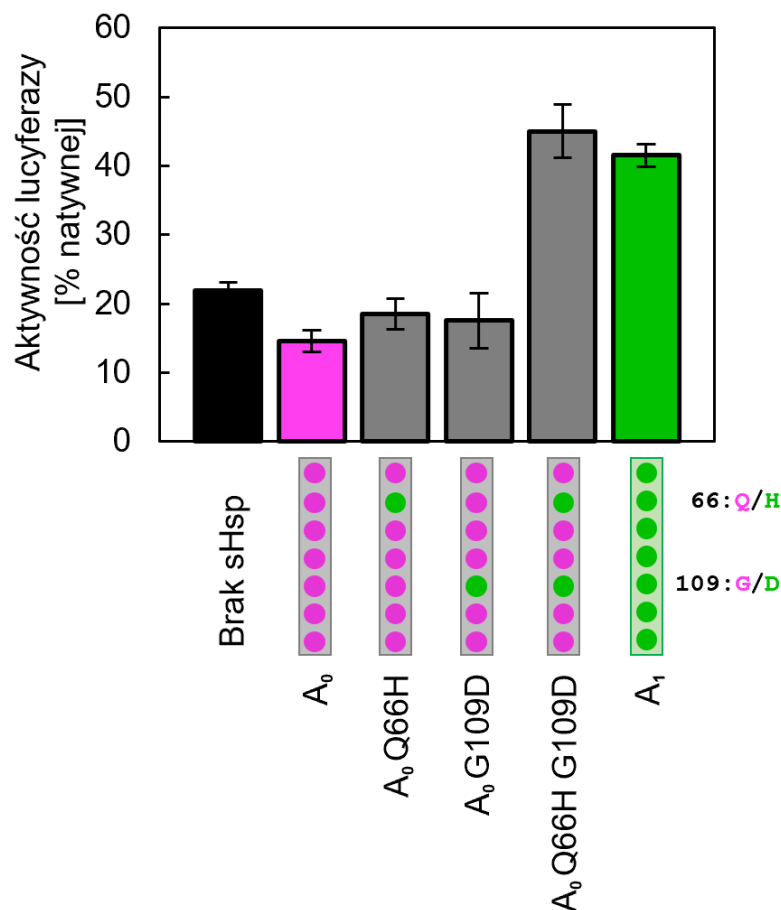
W przypadku wariantów, w których bardziej ancestralny aminokwas znajdował się na pozycji 66 lub 109, zaobserwowałem niemal całkowity brak stymulacji procesu reaktywacji lucyferazy (Ryc. 20). Wskazuje to na fakt, że substytucje Q66H i G109D są wymagane do nabycia zdolności do wydajnej stymulacji reaktywacji lucyferazy przez białko AncA₀.



Ryc. 20 Substytucje na pozycjach 66 i 109 są niezbędne do nabycia przez AncA₀ zdolności do wydajnej stymulacji reaktywacji lucyferazy: Test reaktywacji lucyferazy został przeprowadzony jak uprzednio (ryc. 8); A₀+7 - białko AncA₀ z substytucjami na aminokwas właściwy dla AncA₁ na 7 analizowanych pozycjach; A₀+6* - seria 7 mutantów białka AncA₀, w każdym z których na jednej z 7 analizowanych pozycji znajduje się aminokwas właściwy dla białka AncA₀, a na pozostałych 6 – aminokwas właściwy dla białka AncA₁; substytucje zostały symbolicznie przedstawione na schemacie (ryc. 19).

Aby dowiedzieć się, czy substytucje Q66H i G109D są wystarczające, aby białko AncA₀ uzyskało zdolność do stymulacji reaktywacji lucyferazy podobną do AncA₁, przeprowadziłem analizę trzech kolejnych białek: AncA₀ Q66H, AncA₀ G109D i AncA₀ Q66H G109D.

Białka AncA₀ Q66H i AncA₀ G109D były w stanie stymulować reaktywację lucyferazy w znikomym stopniu, na poziomie porównywalnym z kontrolą nie zawierającą sHsp, podczas gdy białko AncA₀ Q66H G109D wydajnie stymulowało reaktywację lucyferazy, na poziomie zbliżonym do AncA₁ (Ryc. 21). Wynika z tego, że substytucje Q66H i G109D są wystarczające, aby białko AncA₀ nabyło zdolność do wydajnej stymulacji reaktywacji lucyferazy charakterystyczną dla białka AncA₁.

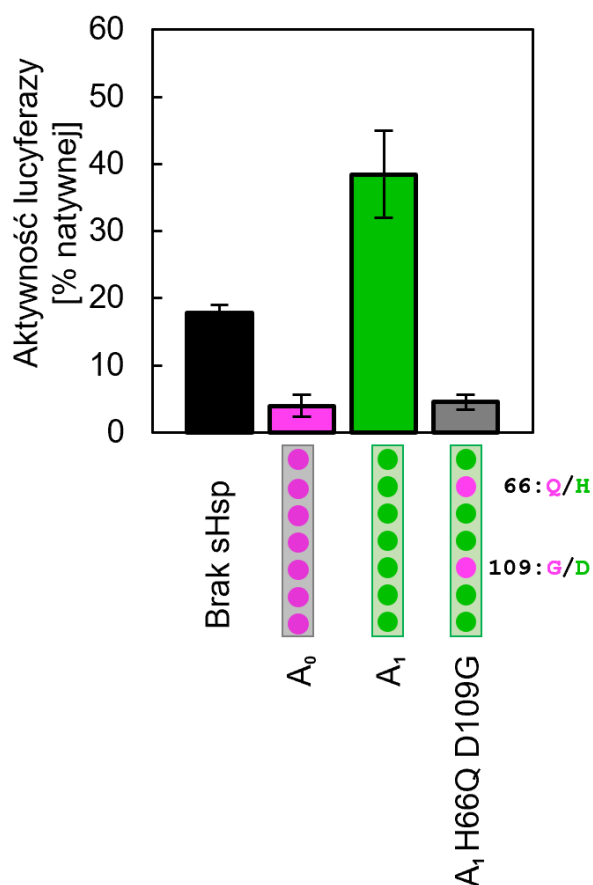


Ryc. 21 Wpływ substytucji na pozycjach 66 i 109 na zdolność białka AncA₀ do stymulacji reaktywacji lucyferazy: Test reaktywacji lucyferazy został przeprowadzony jak uprzednio (ryc. 8); substytucje zostały symbolicznie przedstawione na schemacie (ryc. 19).

Fakt, że białka AncA₀ Q66H i AncA₀ G109D nie były w stanie w znacznym stopniu stymulować reaktywacji lucyferazy wskazuje, że wpływ tych substytucji na efektywność stymulacji dezagregacji i ponownego zwijania zagregowanych białek ma charakter synergiczny – jedynie substytucje na obu tych pozycjach są w stanie doprowadzić do znaczącego zwiększenia zdolności białka AncA₀ do stymulacji tego procesu. Zgadza się to z wcześniejszym wynikiem wskazującym, że substytucja bardziej ancestralnego aminokwasu zarówno na pozycji 66 jak i 109 skutkuje utratą zdolności białka AncA₀ + 7 do efektywnej stymulacji reaktywacji lucyferazy.

Jako dodatkową kontrolę, zbadałem zdolność białka AncA₁ H66Q i D109G (białko AncA₁ z bardziej ancestralnymi aminokwasami na pozycjach 66 i 109) do stymulacji reaktywacji lucyferazy. Wprowadzenie tych dwóch mutacji do białka AncA₁ spowodowało utratę przez nie zdolności do wydajnej stymulacji reaktywacji lucyferazy. Co więcej, białko AncA₁ H66Q D109G inhibowało reaktywację lucyferazy, podobnie do białka AncA₀ (Ryc. 22). Potwierdza to, że substytucje Q66H i G109D

odgrywały kluczową rolę w nabyciu przez białko IbpA w *Erwiniaceae* większej zdolności do stymulacji dezagregacji i ponownego związania białek.

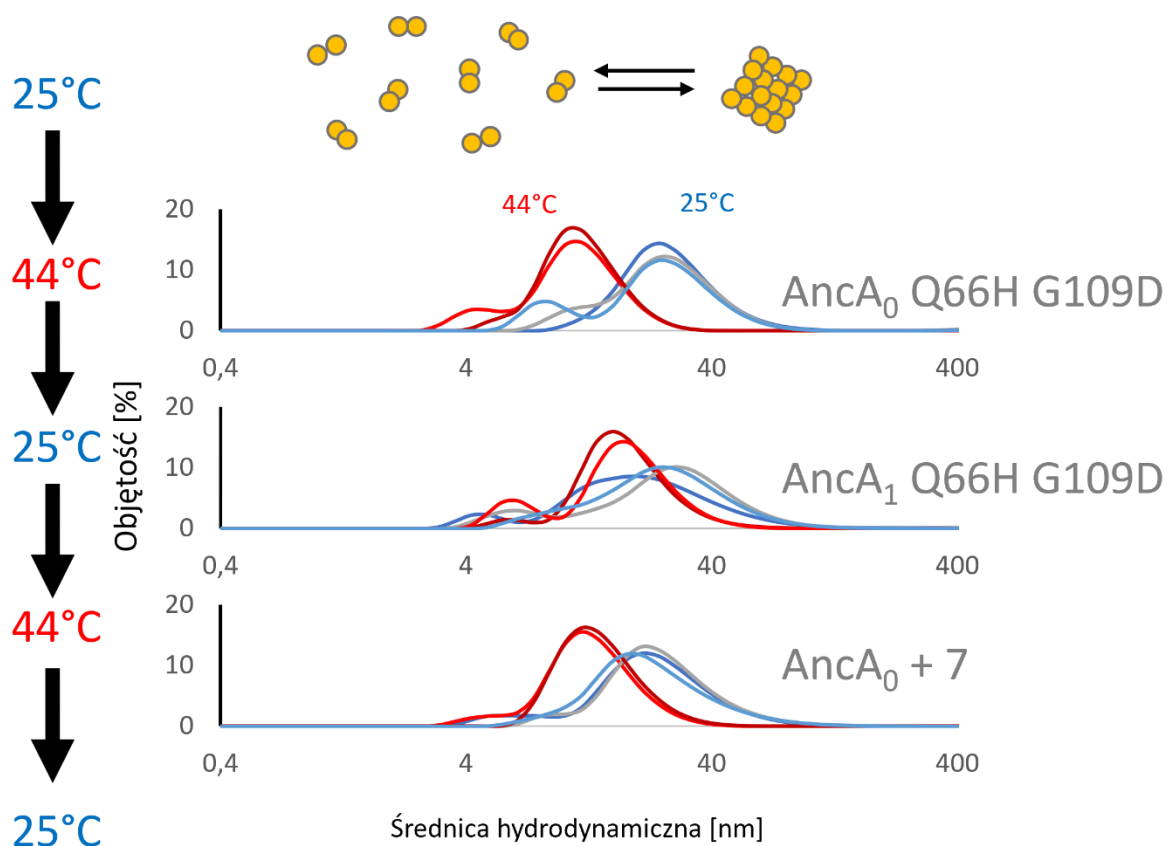


Ryc. 22 Wpływ substytucji na pozycjach 66 i 109 na zdolność białka AncA₁ do stymulacji reaktywacji lucyferazy: Test reaktywacji lucyferazy został przeprowadzony jak uprzednio (ryc. 8); substytucje zostały symbolicznie przedstawione na schemacie (ryc. 19)

Aby zweryfikować, czy wprowadzone substytucje nie zaburzyły pozostałych charakterystycznych funkcji sHsp wykazywanych przez AncA₀ i AncA₁, wykorzystałem technikę DLS w celu zbadania zdolności białek AncA₀ + 7, AncA₀ Q66H G109D i AncA₁ H66Q D109G do odwracalnej deoligomeryzacji w warunkach szoku cieplnego i do sekwestracji agregującego substratu.

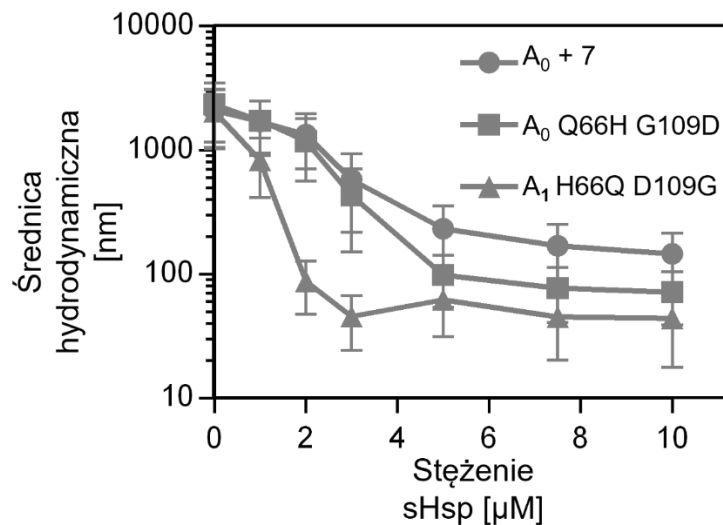
Analizę zdolności do odwracalnej deoligomeryzacji przeprowadziłem jak uprzednio (rozdział 7.4.1). W przypadku każdego z trzech analizowanych białek zaobserwowałem zmniejszenie średnicy hydrodynamicznej tworzonych oligomerów po inkubacji w 44 °C i ponowne jej zwiększenie po inkubacji w 25 °C. Proces ten był w każdym przypadku odwracalny i powtarzalny po ponownej inkubacji w 44 °C i schłodzeniu do 25 °C (Ryc. 23). Wskazuje to, że analizowane mutacje nie zaburzą w

znaczący sposób charakterystycznej dla sHsp odwracalnej deoligomeryzacji w warunkach szoku cieplnego.



Ryc. 23 Białka AncA₀ Q66H G109D, AncA₁ H66Q D109G i AncA₀ + 7 ulegają odwracalnej deoligomeryzacji w warunkach szoku cieplnego: Rozkłady średnic hydrodynamicznych oligomerów tworzonych przez sHsp inkubowane naprzemiennie w 25°C (niebieskie wykresy) i 44°C (czerwone wykresy); średnice hydrodynamiczne wyznaczone za pomocą dynamicznego rozpraszania światła (DLS).

Analizę zdolności do sekwestracji agregującego substratu podobnie przeprowadziłem tak jak uprzednio (rozdział 7.4.3). Każde z analizowanych białek było w stanie, powyżej pewnego stężenia, znacząco zmniejszyć rozmiar cząstek tworzonych przez agregującą w ich obecności lucyferazę. Wskazuje to, że wszystkie posiadały aktywność sekwestrasy. W przypadku białek AncA₀ + 7 i AncA₀ Q66H i G109D, będących w stanie wydajnie stymulować dezagregację (patrz wyżej), wydajność sekwestracji agregującego substratu była niższa, niż w przypadku białka AncA₁ H66Q D109G, będącego inhibitorem dezagregacji (patrz wyżej) (Ryc. 24).



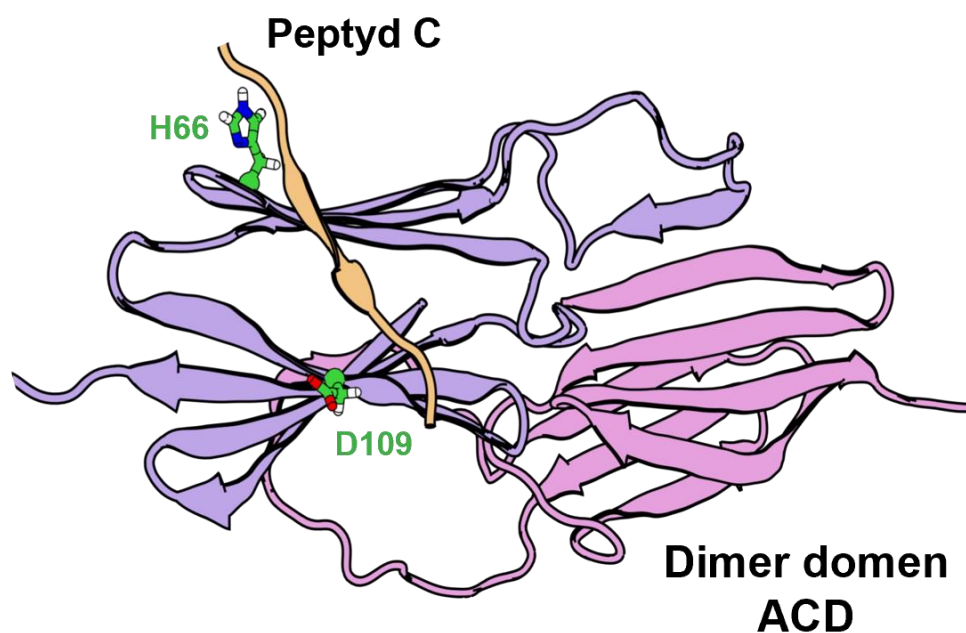
Ryc. 24 Białka AncA₀ Q66H G109D, AncA₁ H66Q D109G i AncA₀ + 7 sekwestrują agregującą lucyferazę: Aktywność sekwestrazy została zmierzona jak uprzednio (ryc. 12).

Wskazuje to, że substytucje na pozycjach 66 i 109 związane są nie tylko ze wzrostem zdolności ancestralnego IbpA do stymulacji dezagregacji i ponownego zwijania substratu, zaobserwowanego pomiędzy węzłami AncA₀ i AncA₁, ale również z towarzyszącym mu spadkiem aktywności sekwestrazy.

7.8 Analiza znaczenia reszt 66 i 109 dla różnic w aktywności białek AncA₀ i AncA₁ na poziomie molekularnym

7.8.1 Lokalizacja reszt 66 i 109 w strukturze białka sHsp

Zidentyfikowane kluczowe reszty 66 i 109 znajdują się w domenie ACD białka IbpA. W celu określenia ich dokładnej lokalizacji w strukturze tego białka, Marcin Pitek, wykorzystując program AlphaFold2, mutagenezę *in silico* i symulację dynamiki molekularnej w stanie równowagi, uzyskał struktury dimerów domen ACD białek AncA₀ i AncA₀ Q66H G109D wraz z zadokowanym peptydem C, reprezentującym region CTE białka AncA₀ (oddziaływanie między tym regionem, a domeną ACD odgrywa istotną rolę w tworzeniu oligomerów sHsp (Stróżecka i in. 2012; Haslbeck, Weinkauff, i Buchner 2019)). Wynika z niej, że reszty 66 i 109 znajdują się w bezpośredniej bliskości kieszeni β4-β8, stanowiącej miejsce wiązania regionu CTE (Ryc. 25).



Ryc. 25 Struktura dimeru domen ACD białka AncA₀ Q66H G109D z zadokowanym peptydem C: Peptyd C reprezentuje region CTE białka AncA₀ o sekwencji aminokwasowej PEAMKPPRIEIN. Aminokwasy na pozycjach 66 i 109 znajdują się w bezpośredniej bliskości kieszeni między wstążkami β4 i β8, stanowiącej miejsce wiązania regionu CTE. Struktura uzyskana przy pomocy programu AlphaFold2 i symulacji dynamiki molekularnej w stanie równowagi w programie Gromacs 2019.2 przez Marcina Pitka.

7.8.2 Analiza wpływu substytucji na pozycjach 66 i 109 na siłę oddziaływania między domeną ACD, a peptydem C *in vitro*

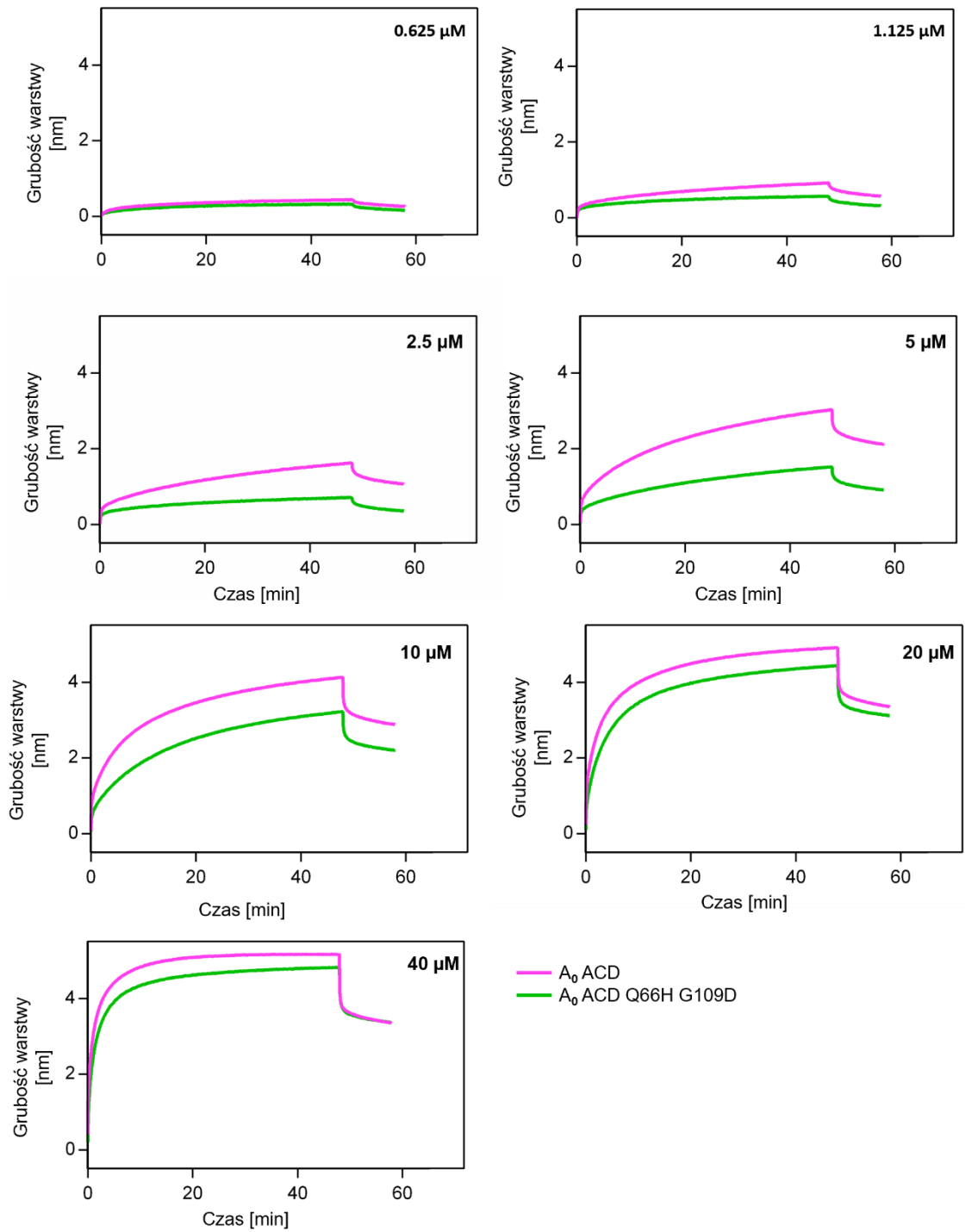
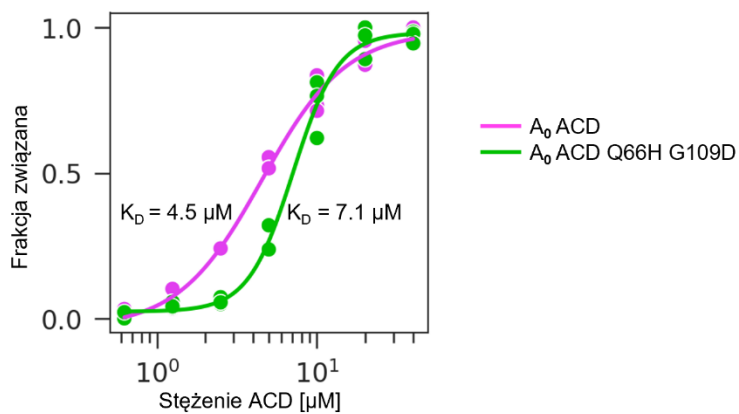
Zidentyfikowane kluczowe reszty 66 i 109 znajdują się w domenie ACD, w bezpośredniej bliskości kieszeni β4-β8. We współczesnych sHsp kieszeń ta jest miejscem wiązania regionu CTE sąsiednich cząsteczek sHsp, które to oddziaływanie jest kluczowe dla funkcjonalności tych białek (Stróżecka i in. 2012; Giese i Vierling 2002).

Aby zweryfikować, czy substytucje na pozycjach 66 i 109, ze względu na ich lokalizację w pobliżu kieszeni β4-β8, wpływają na oddziaływanie między domeną ACD białka AncA₀, a jego regionem CTE, wykorzystałem technikę BLI.

Do analizy wykorzystałem nadprodukowane i oczyszczone domeny ACD białek AncA₀ i AncA₀ Q66H G109D jak i nadprodukowany i oczyszczony peptyd C (odpowiadający regionowi CTE białka AncA₀) z dołączonym łącznikiem SUMO oraz znacznikiem heksahistydynowym. Zastosowanie w analizie samej domeny ACD umożliwiło lepszą kontrolę nad warunkami doświadczenia, niwelując wpływ oligomeryzacji małych białek. Dołączenie łącznika SUMO do peptydu C miało na celu zwiększenie jego rozpuszczalności i zapobieżenie precypitacji w warunkach doświadczenia.

Za pośrednictwem znacznika heksahistydynowego związałem peptyd C do sensora BLI pokrytego jonami niklu. Po usunięciu niezwiązanego peptydu przez płukanie sensora w czystym buforze, obserwowałem wiązanie do sensora badanych domen ACD i ich późniejszą dysocjację. Doświadczenie powtórzyłem dla kolejnych stężeń domen ACD (Ryc. 26 A, B), porównując grubości utworzonej warstwy białek w momencie zakończenia etapu asocjacji dla poszczególnych stężeń obu analizowanych domen.

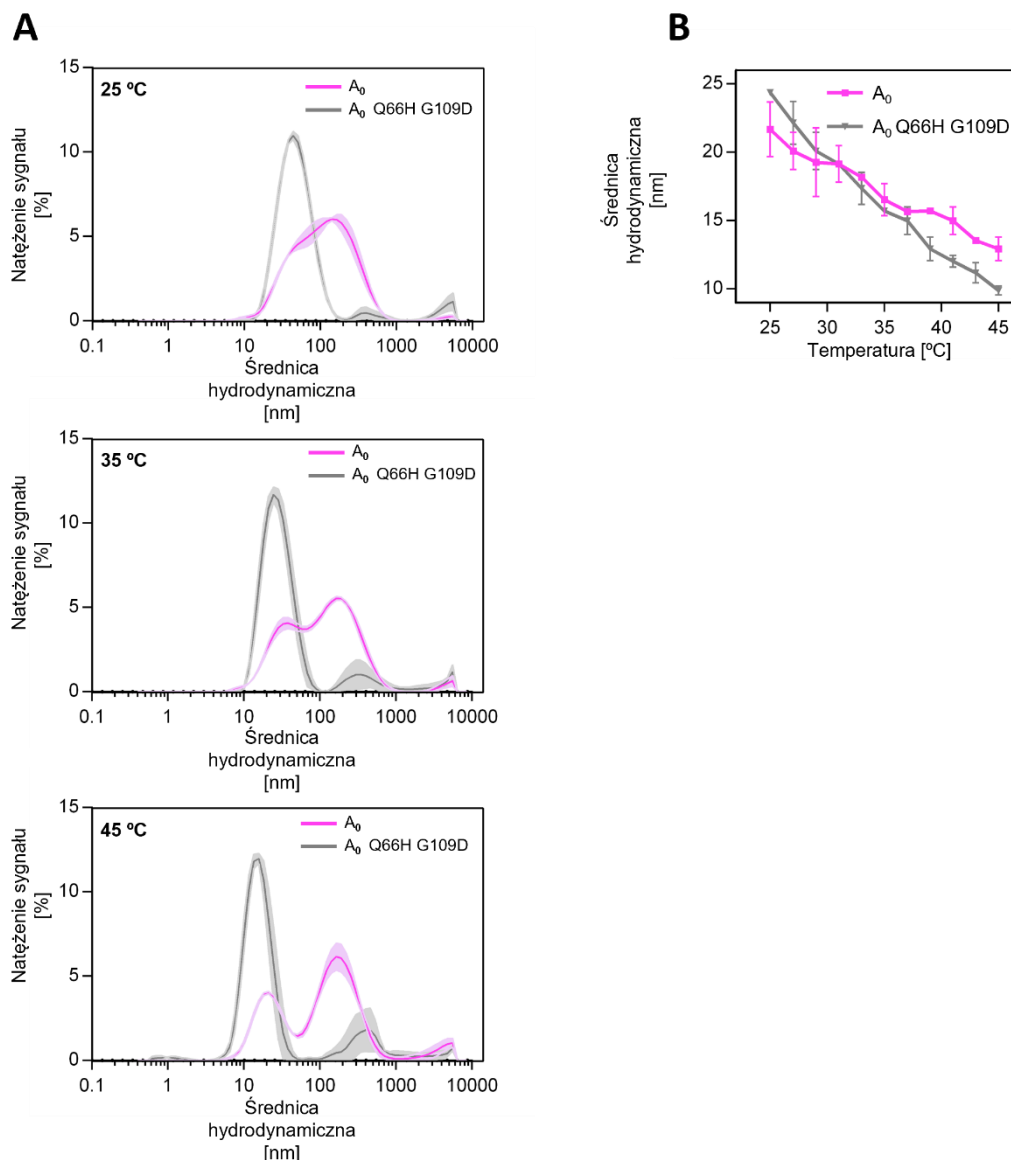
Domena ACD białka AncA₀ charakteryzowała się wyższym powinowactwem do peptydu C, niż domena ACD białka AncA₀ Q66H G109D (Ryc. 26 A, B). Wskazuje to, że substytucje na pozycjach 66 i 109 zmniejszyły w pewnym stopniu powinowactwo między domeną ACD białka AncA₀, a jego regionem CTE.

A**B**

Ryc. 26 Substytucje na pozycjach 66 i 109 zmniejszają powinowactwo domeny ACD białka AncA₀ do regionu CTE: A) Peptyd C został związany do sensora BLI pokrytego jonami niklu za pośrednictwem znacznika Sumo-His₆; sensor był inkubowany przez 45 min z różnymi stężeniami domeny ACD białek AncA₀ i AncA₀ Q66H G109D, po czym został przeniesiony do czystego buforu na 10 min; zmiany grubości warstwy białek na sensorze związane z asocjacją i dysocjacją domeny ACD od peptydu zmierzono przy pomocy techniki BLI; grubość warstwy białek została przedstawiona na wykresie po odjęciu wartości zmierzonych dla próby kontrolnej nie zawierającej domeny ACD. B) Zależność grubości warstwy białek na sensorze w momencie zakończenia etapu asocjacji od użytego stężenia domeny ACD.

7.8.3 Analiza wpływu substytucji na pozycjach 66 i 109 na stan oligomeryczny sHsp

Ze względu na kluczową rolę, jaką odgrywa interakcja między regionem CTE i domeną ACD w tworzeniu oligomerów sHsp (Stróżecka i in. 2012; Giese i Vierling 2002), postanowiłem zbadać, w jaki sposób osłabiające ją substytucje na pozycjach 66 i 109 wpłyną na stan oligomeryczny białka AncA₀. Wykorzystałem w tym celu technikę DLS (rozdział 7.4.1), mierząc rozmiar oligomerów tworzonych przez białka AncA₀ i AncA₀ Q66H G109D w różnych temperaturach. Zaobserwowałem, że substytucje te spowodowały nieznaczne zmniejszenie tworzonych oligomerów jak i pewne ułatwienie deoligomeryzacji zachodzącej na skutek wzrostu temperatury (ryc. 27 A, B).



Ryc. 27 Wpływ substytucji Q66H G109D na stan oligomeryczny białka AncA₀: A) Rozkład średnicy hydrodynamicznej według natężenia sygnału DLS oligomerów tworzonych przez białka AncA₀ i AncA₀ Q66H G109D w różnych temperaturach. B) Zależność temperaturowa średnicy hydrodynamicznej odpowiadającej maksimum dominującego wierzchołka rozkładu średnicy hydrodynamicznej oligomerów AncA₀ i AncA₀ Q66H G109D względem procentu zajmowanej przez objętości.

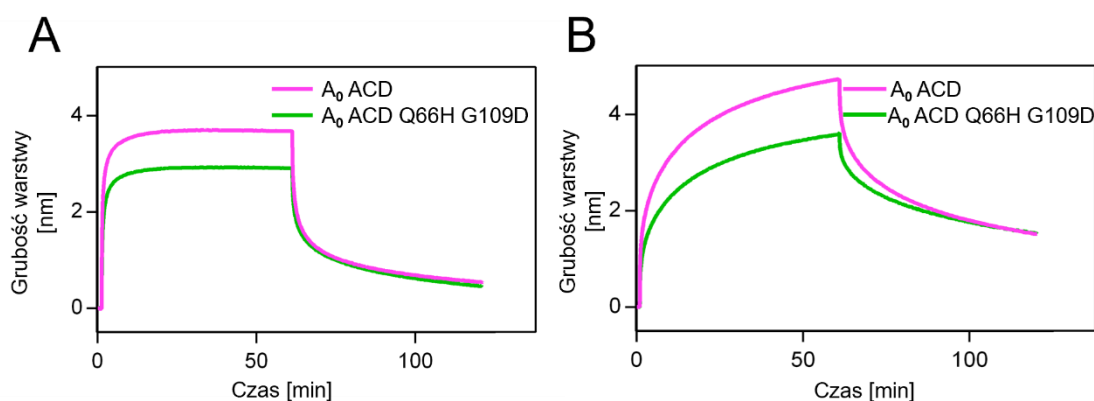
7.8.4 Analiza wpływu substytucji na pozycjach 66 i 109 na wiązanie zagregowanego substratu przez domenę ACD *in vitro*

Wcześniejsze badania wykazały, że zdolność do sekwestracji agregującego substratu i wspomaganie dezagregacji przez sHsp z *E.coli* wymaga obecności regionów terminalnych (Stróżecka i in. 2012). Nie zbadano jednak do tej pory bezpośrednio, czy domena ACD tych białek jest w stanie oddziaływać z substratem.

W przypadku białek sHsp z różnych organizmów zaobserwowano, że domena ACD może oddziaływać z niektórymi substratami (w szczególności amyloidowymi) za pośrednictwem kieszeni między kartkami $\beta 4$ i $\beta 8$ (Jaya, Garcia, i Vierling 2009; Fuchs i in. 2009; Lee i in. 1997; Reinle, Mogk, i Bukau 2022).

Ze względu na to, że ten rejon jest homologiczny do interfejsu zawierającego pozycje 66 i 109 w białku AncA₀, postanowiłem sprawdzić, czy domena ACD tego białka jest w stanie oddziaływać z agregatami białkowymi i czy substytucje na tych dwóch pozycjach modyfikują to oddziaływanie. Podobnie jak w przypadku analizy wiązania całego białka AncA₀ do zagregowanego substratu, wykorzystałem technikę BLI (rozdział 7.4.2). Oprócz zagregowanego lizatu *E. coli* jako dodatkowy modelowy substrat zastosowałem zagregowaną lucyferazę.

W przypadku obu analizowanych substratów zaobserwowałem wiązanie domeny ACD białka AncA₀, jak i AncA₀ Q66H G109D do agregatu (Ryc. 28 A, B). Oznacza to, że domena ACD bakteryjnych sHsp jest w stanie wiązać amorficzne agregaty bez udziału regionów CTE i NTE. Dla obu substratów, domena ACD białka AncA₀ wykazywała się wyższym powinowactwem do substratu niż domena ACD białka AncA₀ Q66H G109D (Ryc. 28 A, B). Wskazuje to, że substytucje na pozycjach 66 i 109 doprowadziły do osłabienia oddziaływania między domeną ACD białka AncA₀, a substratem.

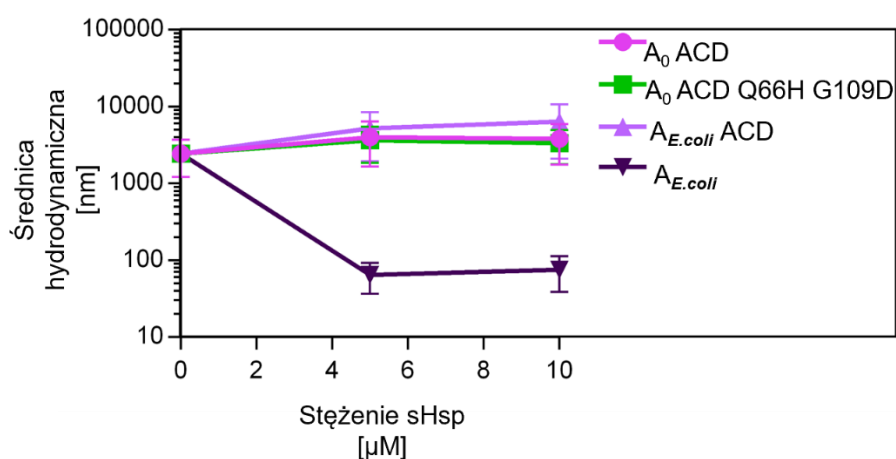


Ryc. 28 Wpływ substytucji Q66H G109D na oddziaływanie domeny ACD białka AncA₀ ze zagregowanym substratem: Wiązanie domeny ACD do zagregowanego substratu zmierzono przy użyciu techniki BLI, jak opisano wyżej (ryc. 11). A) Wiązanie domeny ACD do zagregowanego lizatu *E. coli*. B) Wiązanie domeny ACD do zagregowanej lucyferazy.

Powyższa obserwacja sugeruje również, że zaobserwowane wiązanie między domeną ACD bakteryjnych sHsp i amorficznymi agregatami najprawdopodobniej zachodzi poprzez kieszeń między kartkami $\beta 4$ i $\beta 8$, a potencjalny mechanizm tej interakcji może być zbliżony do mechanizmu interakcji domeny ACD z peptydem C.

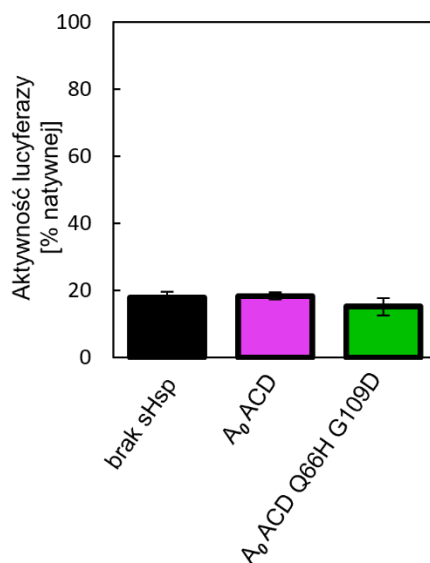
Ze względu na zaobserwowaną zdolność domeny ACD białek AncA₀ i AncA₀ Q66H G109D do wiązania substratu, postanowiłem zbadać, czy posiada ona zdolność do sekwestracji agregujących substratów i wspomagania dezagregacji.

Podobnie jak we wcześniejszych doświadczeniach, do analizy aktywności sekwestrazy wykorzystałem technikę DLS (rozdział 7.4.3). Zarówno w przypadku białka AncA₀, AncA₀ Q66H G109D, jak i współczesnego IbpA z *E. coli*, sama domena ACD nie była w stanie zmniejszyć rozmiaru cząstek tworzonych przez agregującą lucyferazę, niezależnie od stężenia domeny, w przeciwieństwie do pełnego współczesnego białka IbpA z *E. coli* (Ryc. 29). Wynika z tego, że zaobserwowane oddziaływanie między substratem a domeną ACD nie jest wystarczające do wydajnej sekwestracji agregujących substratów.



Ryc. 29 Same domeny ACD nie są w stanie sekwestrować agregującej lucyferazy: Sekwestrację agregującej lucyferazy zmierzono jak opisano wyżej (ryc. 12). jako kontrolę wykorzystano pełnej długości współczesne białko IbpA z *E. coli*

Do analizy stymulacji dezagregacji wykorzystałem, jak uprzednio, test reaktywacji lucyferazy (rozdział 7.4.4). Sama domena ACD zarówno w przypadku białka AncA₀, jak i AncA₀ Q66H G109D, nie była w stanie stymulować ani hamować reaktywacji lucyferazy przez system Hsp70 i Hsp100 (Ryc. 30). Wskazuje to, że, podobnie jak w przypadku aktywności sekwestrazy, samo zaobserwowane oddziaływanie między domeną ACD, a substratem nie wystarcza do stymulacji dezagregacji.



Ryc. 30 Same domeny ACD nie są w stanie stymulować reaktywacji lucyferazy: Test reaktywacji lucyferazy został przeprowadzony jak opisano wyżej (ryc. 8)

Potwierdza to wyniki wcześniejszych badań wykazujące, że regiony NTE i CTE są niezbędne dla aktywności bakteryjnych sHsp (Stróżecka i in. 2012), zarówno w przypadku sekwestracji agregujących substratów jak i stymulacji dezagregacji.

Powyższe wyniki wskazują, że substytucje Q66H i G109D doprowadziły do osłabienia oddziaływania między domeną ACD białka IbpA, a regionem CTE, jak i między tą domeną, a substratem w toku ewolucji wtórnie pojedynczego IbpA w *Erwiniaceae*.

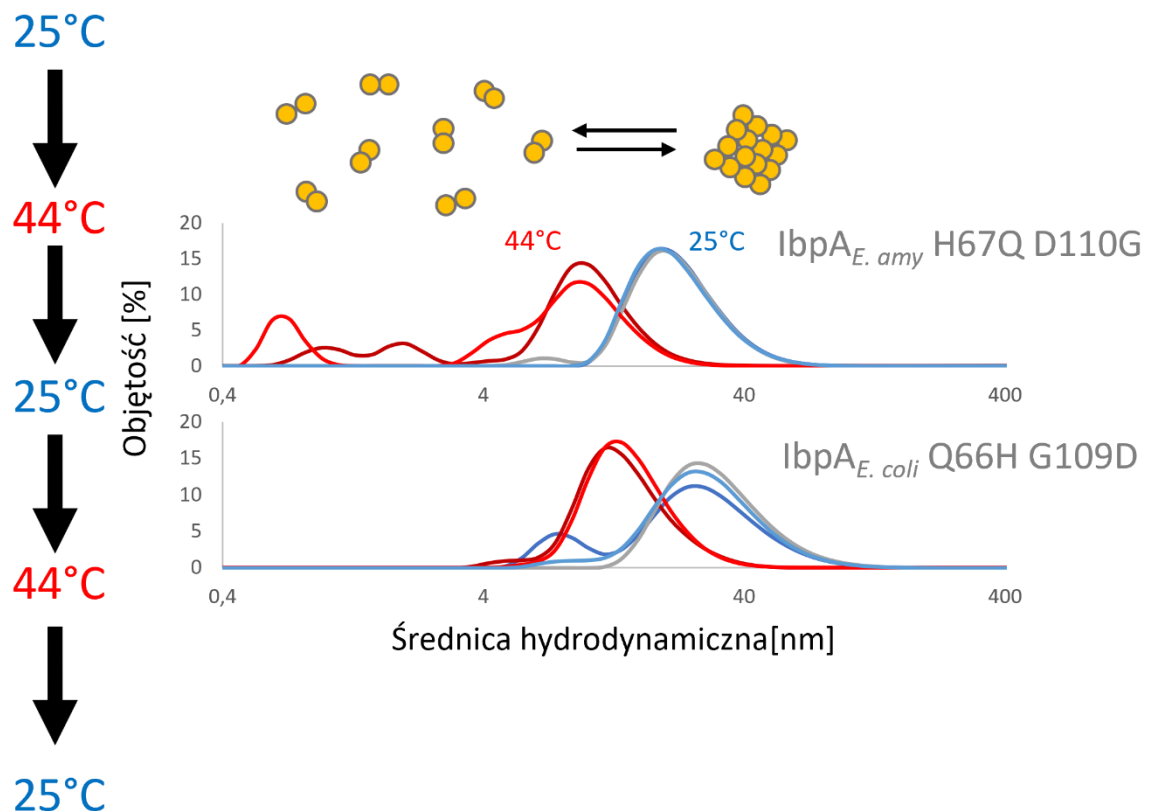
Zmiany te mogą tłumaczyć obserwowany spadek aktywności sekwestrasy i wzrost zdolności do stymulacji dezagregacji – osłabienie oddziaływania z substratem jak i potencjalnie oddziaływania pomiędzy poszczególnymi sHsp w kompleksie z substratem (poprzez oddziaływanie regionów CTE z domenami ACD sąsiednich białek) mogło ułatwić konkurencję systemu Hsp70 o miejsca wiązania na substracie w trakcie dezagregacji, jednocześnie w pewnym stopniu utrudniając wiązanie substratu przez sHsp w trakcie sekwestracji.

7.9 Analiza znaczenia pozycji 66 i 109 we współczesnych białkach IbpA z *E. coli* i *E. amylovora*

Kolejnym krokiem było zbadanie, jaką rolę odgrywają pozycje 66 i 109 lub homologiczne w determinacji różnic funkcjonalnych między współczesnymi białkami IbpA z *Erwiniaceae* i *Enterobacteriaceae*. Aminokwasy występujące na tych pozycjach w IbpA z *E. coli* są identyczne jak w przypadku chronologicznie starszego, rekonstruowanego białka AncA₀, podczas gdy aminokwasy występujące na homologicznych pozycjach (67 i 110) w IbpA z *E. amylovora* są identyczne jak w przypadku chronologicznie nowszego, rekonstruowanego białka AncA₁.

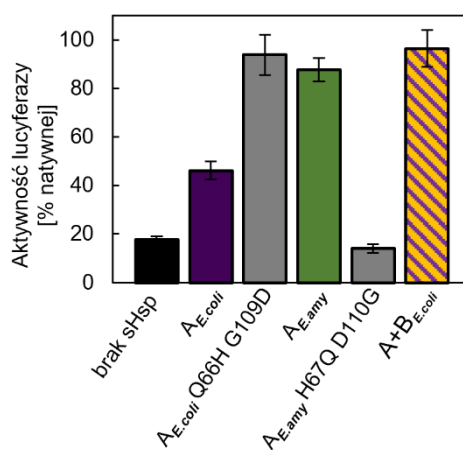
Aby zbadać, czy te różnice wystarczają do wyjaśnienia różnic funkcjonalnych między tymi współczesnymi białkami, sprawdziłem, jak wprowadzenie aminokwasów charakterystycznych dla IbpA z *E. amylovora* na pozycjach 66 i 109 w IbpA z *E. coli* oraz aminokwasów charakterystycznych dla IbpA z *E. coli* na homologicznych pozycjach 67 i 110 w IbpA z *E. amylovora* wpływa na ich zdolność tych białek wspomaganą dezagregacji i ponownego zwijania substratu, wiązania substratu i sekwestracji agregującego substratu.

Przed właściwymi analizami zweryfikowałem, czy wprowadzone substytucje nie zaburzają charakterystycznej dla sHsp odwracalnej deoligomeryzacji w warunkach szoku cieplnego, wykorzystując jak uprzednio technikę DLS (rozdział 6.4.1). Zarówno w przypadku białka IbpA_{*E. coli*} Q66H G109D, jak i IbpA_{*E. amylovora*} H67Q D110G tworzone przez nie oligomery uległy zmniejszeniu po inkubacji w 44 °C i powiększeniu po ponownej inkubacji w 25 °C. Efekt ten był odwracalny i powtarzalny po ponownej inkubacji w 44 i 25 °C (Ryc.31). Wynika z tego, że wprowadzone substytucje nie doprowadziły do utraty zdolności białek IbpA z *E. coli* i *E. amylovora* do odwracalnej deoligomeryzacji w warunkach szoku cieplnego.



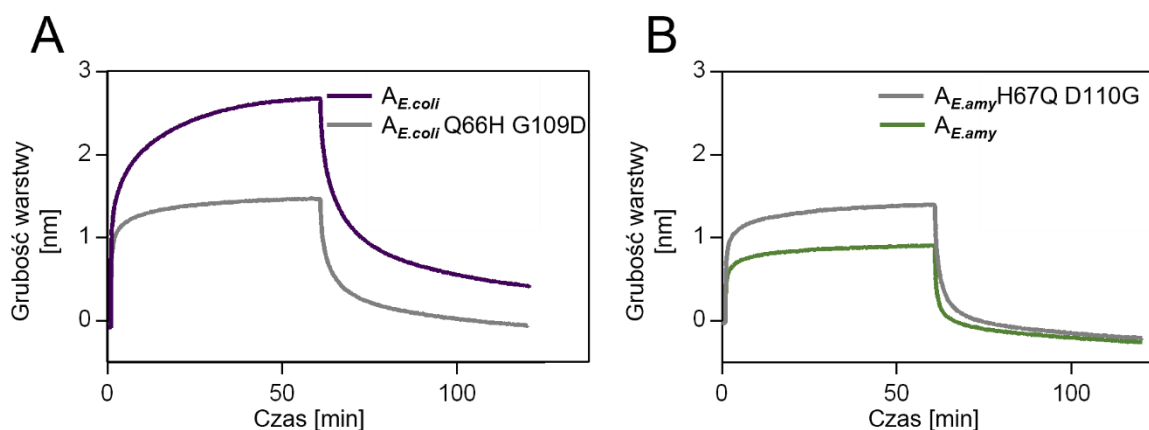
Ryc. 31 Współczesne białka IbpA z *E. coli* i *E. amylovora* z substytucjami na pozycjach homologicznych do pozycji 66 i 109 w AncA₀ odwracalnie deoligomeryzują w warunkach szoku cieplnego: Rozkłady średnic hydrodynamicznych oligomerów tworzonych przez sHsp inkubowane naprzemiennie w 25°C (niebieskie wykresy) i 44°C (czerwone wykresy); średnice hydrodynamiczne wyznaczone za pomocą dynamicznego rozpraszania światła (DLS).

W celu analizy stymulacji dezagregacji i ponownego zwijania substratów, tak jak w poprzednich przypadkach wykorzystałem test reaktywacji lucyferazy (rozdział 7.4.4). Białko lbpA_{E.coli} było w stanie jedynie w niewielkim stopniu stymulować reaktywację lucyferazy, podczas gdy białko lbpA_{E.coli} Q66H G109D było w stanie stymulować reaktywację lucyferazy niemal dwa razy wydajniej, podobnie jak białko lbpA_{E.amylovora}. Białko lbpA_{E.amylovora} H67Q D110G z kolei utraciło zdolność do wydajnej stymulacji reaktywacji lucyferazy w porównaniu z typem dzikim (Ryc. 32). Wynika z tego, że różnice na badanych pozycjach są przyczyną większej zdolności białka lbpA z *E. amylovora* do stymulacji dezagregacji i ponownego zwijania substratów w porównaniu z białkiem lbpA z *E.coli*.



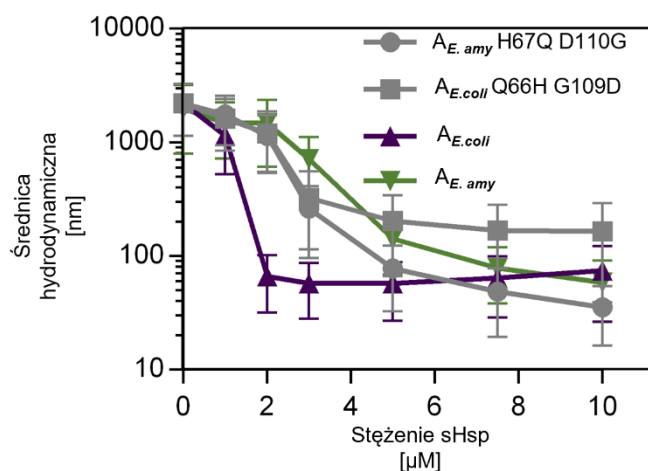
Ryc. 32 Wpływ substytucji na pozycjach homologicznych do pozycji 66 i 109 w AncA₀ na zdolność współczesnych białek lbpA z *E. coli* i *E. amylovora* do stymulacji reaktywacji lucyferazy: Test reaktywacji lucyferazy został przeprowadzony jak opisano wyżej (ryc. 8)

Do analizy wpływu substytucji na pozycjach 66 i 109 na powinowactwo współczesnych białek lbpA do substratu wykorzystałem technikę BLI (rozdział 7.4.3). Wprowadzenie substytucji Q66H G109D doprowadziło do zmniejszenia powinowactwa lbpA z *E. coli* do substratu (Ryc. 33 A), podczas gdy wprowadzenie substytucji H67Q i D110G doprowadziło do wzrostu powinowactwa lbpA z *E. amylovora* do substratu (Ryc. 33 B). Świadczy to o tym, że różnice na badanych pozycjach są przyczyną mniejszego powinowactwa lbpA z *E. amylovora* do substratu w porównaniu z lbpA z *E. coli*.



Ryc. 33 Wpływ substytucji na pozycjach homologicznych do pozycji 66 i 109 w AncA₀ na wiązanie współczesnych białek IbpA z *E. coli* i *E. amylovora* do zagregowanej lucyferazy: Wiązanie współczesnych sHsp typu dzikiego i mutantów z *E. coli* (A) i *E. amylovora* (B) do zagregowanej lucyferazy zmierzono przy użyciu techniki BLI jak opisano powyżej (ryc. 11).

Wpływ substytucji na badanych pozycjach na aktywność sekwestrazy białek IbpA z *E. coli* i *E. amylovora* zbadałem przy użyciu techniki DLS (rozdział 7.4.3). W przypadku współczesnego białka IbpA z *E.coli*, wprowadzenie substytucji Q66H G109D doprowadziło do znaczącego zmniejszenia zdolności do wydajnej sekwestracji agregującego substratu, podczas gdy w przypadku współczesnego białka IbpA z *E. amylovora* wprowadzenie substytucji H67Q D110G doprowadziło do wzrostu zdolności do wydajnej sekwestracji agregującego substratu (Ryc. 34). Wynika z tego, że różnice na badanych pozycjach przynajmniej częściowo odpowiadają za różnice w aktywności sekwestrazy między współczesnymi białkami IbpA z *E. coli* i *E. amylovora*.

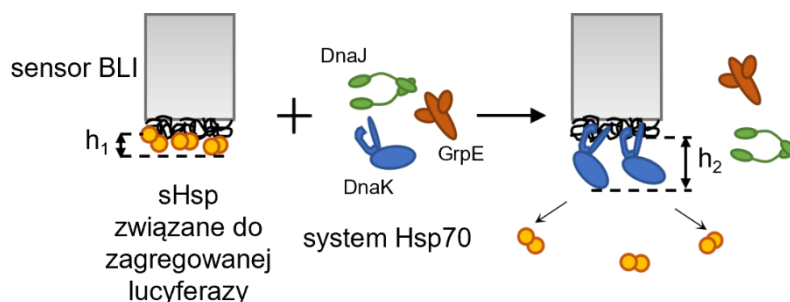


Ryc. 34 Wpływ substytucji na pozycjach homologicznych do pozycji 66 i 109 w AncA₀ na zdolność współczesnych białek IbpA z *E. coli* i *E. amylovora* do sekwestracji agregującej lucyferazy: Aktywność sekwestrazy badanych sHsp została zmierzona jak opisano powyżej (ryc. 12);

Powyższe wyniki wskazują, że różnice aminokwasów na badanych dwóch pozycjach w znacznej mierze odpowiadają za obserwowane różnice między współczesnymi białkami IbpA z *E. coli* i *E. amylovora* w aktywności sekwestrazy i zdolności do stymulacji dezagregacji i ponownego związania białek. Na poziomie molekularnym najprawdopodobniej jest to związane, jak w przypadku białek ancestralnych, z różnicami w powinowactwie domeny ACD do substratu i regionu CTE, co może wpływać na konkurencję między sHsp, a systemem Hsp70 o miejsca wiązania na substracie w procesie dezagregacji (rozdział 7.8).

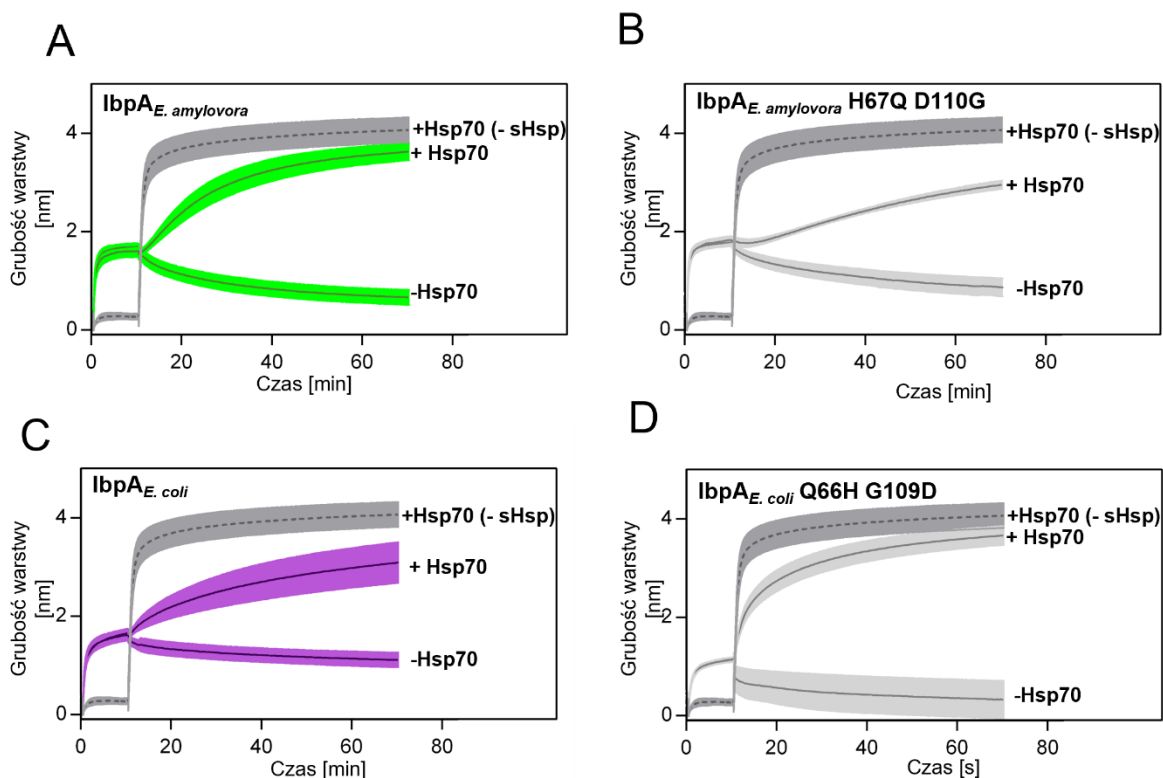
Aby zweryfikować powyższą hipotezę, zbadałem, w jakim stopniu substytucje na badanych pozycjach wpływają na inhibicję wiązania systemu Hsp70 do agregatów przez białka IbpA z *E. coli* i *E. amylovora*. W tym celu wykorzystałem technikę BLI (rozdział 7.4.2).

Sensor BLI ze związanym agregatem lucyferazy inkubowałem w roztworze sHsp w celu opłaszczenia nimi agregatu, a następnie przenieśliem do buforu zawierającego system Hsp70 (DnaK, DnaJ, GrpE) i ATP. Ze względu na fakt, że system Hsp70, wiążąc się do agregatu w warunkach doświadczenia, tworzył warstwę o grubości znacznie większej niż utworzona przez sHsp, byłem w stanie zaobserwować zastępowanie na agregacie sHsp przez system Hsp70 (Ryc. 35). Porównując uzyskane wyniki z kontrolą, w której sensor pokryty agregatem białkowym był inkubowany z systemem Hsp70 pod nieobecność sHsp jak i kontrolą, w której sHsp dysocjowały od agregatu pod nieobecność systemu Hsp70, byłem w stanie określić, w jakim stopniu obecność danego sHsp hamowała wiązanie systemu Hsp70 do substratu.



Ryc. 35 Schemat pomiaru wypierania sHsp z agregatu przez białka systemu Hsp70: Sensor BLI ze związaną zagregowaną lucyferazą był inkubowany w obecności $5 \mu\text{M}$ aktywowanych termicznie sHsp przez 10 min w celu opłaszczenia przez nie agregatu. Tak przygotowany sensor został następnie przeniesiony do roztworu zawierającego system Hsp70 ($0,7 \mu\text{M}$ DnaK, $0,28 \mu\text{M}$ DnaJ, $0,21 \mu\text{M}$ GrpE) i inkubowany przez 60 min. Ponieważ grubość warstwy utworzonej na agregacie lucyferazy przez system Hsp70 jest większa od grubości warstwy tworzonej przez sHsp ($h_2 > h_1$), zastępowanie sHsp na agregacie przez system Hsp70 widoczne jest jako przyrost grubości warstwy białek podczas pomiaru BLI. Porównanie z kontrolą, w której system Hsp70 jest inkubowany z sensorem ze związaną zagregowaną lucyferazą pod nieobecność sHsp pozwala na określenie, w jakim stopniu sHsp utrudniają wiązanie systemu Hsp70 do agregatu.

Wszystkie z badanych sHsp w pewnym stopniu hamowały wiązanie systemu Hsp70 do agregatów lucyferazy. W przypadku współczesnego IbpA z *E. coli* (Ryc. 36 C) inhibicja ta była silniejsza niż w przypadku białka IbpA z *E. amylovora* (Ryc. 36 A). Wprowadzenie substytucji Q66H G109D do IbpA z *E. coli* spowodowało osłabienie tej inhibicji do poziomu porównywalnego z białkiem z *E. amylovora* (Ryc. 36 D), podczas gdy wprowadzenie substytucji H67Q D110G do białka IbpA z *E. amylovora* zwiększyło obserwowaną inhibicję do poziomu porównywalnego z białkiem IbpA z *E. coli* (Ryc. 36 B). Wynika z tego, że różnice w aminokwasach obecnych na pozycjach 66 i 109 w IbpA z *E. coli* i homologicznych pozycjach 67 i 110 w IbpA z *E. amylovora* sprawiają, że białka te w różnym stopniu hamują wiązanie systemu Hsp70 do substratów związanych w kompleksie z sHsp.



Ryc. 36 Wpływ substytucji na pozycjach homologicznych do pozycji 66 i 109 w AncA₀ na zdolność współczesnych białek IbpA z *E. coli* i *E. amylovora* do inhibicji wiązania systemu Hsp70 do substratu: Doświadczenie przeprowadzono jak opisano powyżej (Ryc. 35); na każdym wykresie przedstawiono właściwe doświadczenie – wiązanie systemu Hsp70 do zagregowanej lucyferazy opłaszczanej sHsp (+Hsp70), jak i dwie kontrole: wiązanie systemu Hsp70 do zagregowanej lucyferazy pod nieobecność sHsp (+Hsp70 (- sHsp)) oraz dysocjację sHsp od zagregowanej lucyferazy pod nieobecność systemu Hsp70 (-sHsp) Wykresy przedstawiono jako średnią grubość warstwy białek ± odchylenie standardowe. A) Współczesne IbpA z *E. amylovora*. B) Współczesne IbpA z *E. amylovora* z wprowadzonymi substytucjami H67Q D110G. C) Współczesne IbpA z *E. coli*. D) Współczesne IbpA z *E. coli* z wprowadzonymi substytucjami Q66H G109D.

Analiza substytucji na pozycjach 66 i 109 we współczesnym białku IbpA z *E. coli* jak i na homologicznych pozycjach 67 i 110 we współczesnym białku IbpA z *E. amylovora* pokazuje, że aminokwasy znajdujące się na tych pozycjach w istotnym stopniu odpowiadają za różnice funkcjonalne między tymi białkami pod względem ich zdolności do wiązania i sekwestracji substratów, stymulacji prowadzonej przez system Hsp70 i białka Hsp100 dezagregacji i przywracania konformacji natywnej sekwestrowanych substratów jak i łatwości, z jaką system Hsp70 może konkurować z nimi o miejsca wiązania na substracie.

8. Dyskusja

W ramach niniejszej pracy przeprowadziłem analizę historii ewolucyjnej białek IbpA w rzędzie *Enterobacterales* w celu określenia, w jaki sposób wtórnie pojedyncze białko IbpA z rodziny *Erwiniaceae* nabyło zdolność do efektywnego wspomaganie dezagregacji prowadzonej przez system Hsp70 i białka Hsp100 bez współpracy z białkiem IbpB.

Zrekonstruowane drzewo filogenetyczne, opisujące pokrewieństwo pomiędzy ortologami białka IbpA w rzędzie *Enterobacterales* (rozdział 7.2) cechuje się wysokim podobieństwem do drzew filogenetycznych otrzymanych na podstawie sekwencji wielu białek, m. in. z wyraźnym podziałem rodziny *Erwiniaceae* na rodzaje *Erwinia*, *Pantoea*, *Tatumella* i *Mixta* (Obuchowski i in. 2019; Palmer i in. 2018). Przeprowadzona analiza bootstrap wykazuje wysokie wsparcie dla kluczowych węzłów filogenezy, co wspiera wiarygodność przeprowadzonej rekonstrukcji. Sugeruje to również, że, przynajmniej na poziomie rodzin, horyzontalny transfer genów, bądź niekompletne sortowanie się alleli odgrywały niewielką rolę w ewolucji białek sHsp w rzędzie *Enterobacterales*.

Weryfikacja zgodności drzewa filogenetycznego białek IbpA z drzewem filogenetycznym gatunków jest trudniejsza w obrębie poszczególnych rodzajów. Klasyfikacja taksonomiczna pojedynczych gatunków do rodzajów nie zawsze w pełni odwzorowuje ich rzeczywiste pokrewieństwo ewolucyjne. Może to częściowo tłumaczyć fakt, że na otrzymanym drzewie filogenetycznym pojedyncze sekwencje białek pochodzących z organizmów klasyfikowanych jako należące do rodzaju *Pantoea* grupowały się z rodzajem *Erwinia* i *vice versa*. Podobna sytuacja może mieć miejsce w przypadku rodzajów *Pantoea* i *Mixta*, które zostały sklasyfikowane jako odrębne rodzaje stosunkowo niedawno (Palmer i in. 2018).

Analiza właściwości rekonstruowanych białek ancestralnych pokazuje, że zdolność wtórnie pojedynczego białka IbpA do wydajnego wspomaganie dezagregacji i ponownego zwijania sekwestrowanych substratów bez współpracy z białkiem IbpB najprawdopodobniej wykształciła się już u ostatniego wspólnego przodka rodziny *Erwiniaceae*, równoległe z utratą IbpB (rozdział 7.3, 7.4, 7.5). Nasuwa to pytanie o zależność przyczynowo – skutkową: czy nabycie nowej funkcji przez białko IbpA umożliwiło utratę białka IbpB, czy też utrata białka IbpB stała się źródłem presji selekcyjnej powodującej nabycie nowej funkcji przez białko IbpA?

Analiza selekcji naturalnej oparta o stosunek liczby substytucji niesynonimicznych i synonimicznych (rozdział 7.6) pokazuje, że nabyciu nowej funkcjonalności przez wtórnie pojedyncze IbpA w *Erwiniaceae* z wysokim prawdopodobieństwem towarzyszył epizod selekcji pozytywnej działającej na to białko. Sugeruje to, że zaobserwowane zmiany miały prawdopodobnie charakter

adaptacyjny, a hipoteza, że białko IbpA nabyło nową funkcjonalność na skutek dryfu genetycznego, co umożliwiło utratę białka IbpB, jest mało prawdopodobna.

Istnieje możliwość, że ancestralny dwubiałkowy system sHsp w ostatnim wspólnym przodku rodzin *Erwiniaceae* i *Enterobacteriaceae* cechował się mniejszą efektywnością niż współczesne systemy dwubiałkowe, tak że nowa funkcjonalność białka IbpA prowadziła do uzyskania bardziej efektywnego systemu i była promowana przez selekcję pozytywną. Nowa funkcjonalność IbpA umożliwiła następnie utratę białka IbpB dzięki rozluźnieniu presji selekcyjnej na utrzymanie systemu dwubiałkowego.

Nie można jednak całkowicie wykluczyć hipotezy, że ostatni wspólny przodek *Erwiniaceae* znajdował się pod presją selekcyjną promującą system jednobiałkowy (gdy korzyści z mniejszego genomu i mniejszej złożoności proteomu przeważały nad korzyściami z wyższej efektywności systemu dwubiałkowego w przywracaniu struktury natywnej sekwencjonowanych białek). Mogło to doprowadzić do utraty białka IbpB, co stało się następnie źródłem presji selekcyjnej promującej nową funkcjonalność wtórnie pojedynczego IbpA.

Weryfikację obu zaproponowanych hipotez utrudnia brak wiedzy na temat właściwości ancestralnego białka IbpB, będącego partnerem białka AncA₀ w ostatnim wspólnym przodku rodzin *Erwiniaceae* i *Enterobacteriaceae*. Ze względu na brak ortologów białka IbpB we współczesnych *Erwiniaceae*, niemożliwa jest rekonstrukcja ancestralna tego konkretnego białka. Istnieje jednak możliwość, że rekonstrukcja ancestralna systemów dwubiałkowych z poprzedzającego i kolejnego węzła drzewa filogenetycznego pomogłaby wnioskować na temat dokładnych właściwości dwubiałkowego systemu sHsp w ostatnim wspólnym przodku rodzin *Erwiniaceae* i *Enterobacteriaceae* i precyzyjnych przyczyn utraty białka IbpB.

Analiza alternatywnych rekonstruowanych białek ancestralnych, które na każdej niepewnej pozycji posiadały resztę aminokwasową o drugim najwyższym prawdopodobieństwie *a posteriori*, pokazała, że obserwowane właściwości rekonstruowanych białek nie zależą od niepewności statystycznej związanej z rekonstrukcją ancestralną.

Porównanie sekwencji rekonstruowanych białek AncA₀ i AncA₁ w połączeniu z analizą ich mutantów pozwoliło mi zidentyfikować dwie substytucje aminokwasowe (Q66H i G109D), zlokalizowane w obszarze kieszeni β 4- β 8 domeny ACD, które odegrały kluczową rolę w nabyciu nowej funkcjonalności przez wtórnie pojedyncze białko IbpA z *Erwiniaceae* (rozdział 7.7). Substytucje te posiadały efekt synergistyczny, tak że każda z nich pojedynczo nie była wystarczająca, aby nadać bardziej ancestralnemu białku AncA₀ funkcjonalność typową dla wtórnie pojedynczego IbpA z *Erwiniaceae*.

Wprowadzenie zidentyfikowanych substytucji na homologicznych pozycjach we współczesnym białku IbpA z *Escherichia coli* spowodowało osłabienie aktywności sekwestrazy i zdolności do wiązania zagregowanego substratu przy jednoczesnym zwiększeniu jego zdolności do stymulacji dezagregacji i ponownego zwijania substratów. Odwrócenie tych substytucji na homologicznych pozycjach w białku IbpA z *Erwinia amylovora* sprawiło natomiast, że jego aktywność sekwestrazy i zdolność wiązania zagregowanego substratu wzrosła, a zdolność do stymulacji dezagregacji i ponownego zwijania substratu zmalała (rozdział 7.9). Potwierdza to, że zidentyfikowane substytucje determinują różnice funkcjonalne między współczesnymi ortologami IbpA z rodzin *Erwiniaceae* i *Enterobacteriaceae*.

Zdarza się, że substytucje zidentyfikowane w procesie rekonstrukcji ancestralnej jako kluczowe dla nabycia nowej funkcji nie posiadają analogicznego efektu po wprowadzeniu do białek współczesnych, ze względu na interakcje epistatyczne i zależność od szerszego kontekstu sekwencji (Natarajan i in. 2023). Nie zaobserwowałem tego w przypadku zidentyfikowanych przeze mnie substytucji, co sugeruje, że ich funkcja w ortologach IbpA w *Erwiniaceae* i *Enterobacteriaceae* nie jest silnie zależna od szerszego kontekstu sekwencji aminokwasowej.

Rekonstruowane białko AncA₁, mimo iż było w stanie wyraźnie efektywniej stymulować dezagregację i przywracanie konformacji natywnej substratów niż białko AncA₀, wciąż charakteryzowało się mniejszą efektywnością niż współczesne białko z *E. amylovora*. Podobnie współczesne białko IbpA z *E. coli* wykazuje się zdolnością do słabej stymulacji reaktywacji lucyferazy w warunkach, w których białko AncA₀ jest inhibitorem tego procesu. Sugeruje to, że jakkolwiek dwie zidentyfikowane substytucje odgrywają kluczową rolę w różnicowaniu aktywności między białkami IbpA z systemu dwubiałkowego i wtórnie pojedynczymi białkami IbpA, substytucje na innych pozycjach w sHsp, które zaszły w toku ewolucji zarówno w rodzinie *Enterobacteriaceae* jak i *Erwiniaceae* również miały pewien wpływ na tą aktywność.

Należy zwrócić uwagę, że pozycja 66 jest w białku AncA₀ rekonstruowana z wysoką niepewnością (rozdział 7.5) – z prawdopodobieństwem *a posteriori* 30% w białku tym na pozycji 66 mogła znajdować się histydyna zamiast glutaminy, a substytucja Q66H mogła nastąpić już na poprzedzającej gałęzi drzewa filogenetycznego. Taki scenariusz nie wpływa jednak na wniosek, że nabycie nowej funkcjonalności wtórnie pojedynczego IbpA nastąpiło dopiero między węzłami AncA₀ i AncA₁ – ze względu na zaobserwowaną synergię jedynie zajście obu tych substytucji umożliwia wtórnie pojedynczemu białku IbpA efektywną stymulację dezagregacji i ponownego zwijania substratów pod nieobecność IbpB. Potwierdza to obserwacja, że białko AncA₀Alt, reprezentujące alternatywny wariant rekonstrukcji ancestralnego białka AncA₀, posiada właściwości zbliżone do AncA₀ mimo posiadania histydyny na pozycji 66 (rozdział 7.5).

Co więcej, jakkolwiek niemal wszystkie współczesne białka IbpA wchodzące w skład systemu dwubiałkowego posiadają glicynę na pozycji 109 (lub homologicznej), białka IbpA pochodzące z rodzajów *Chania* i *Serratia* oraz rodziny *Hafniaceae* (numery akcesji w bazie UniProt: C5BHD3, A0A097QX83, W0LEB5, L0MCB7, A8G7R8) posiadają histydynę na pozycji 66 (lub homologicznej). Sugeruje to, że substytucja Q66H mogła zajść kilkakrotnie w historii ewolucyjnej białek IbpA, ale samodzielnie miała ona prawdopodobnie charakter neutralny lub zbliżony do neutralnego. Dopiero zajście dodatkowej substytucji G109D w ostatnim wspólnym przodku rodziny *Erwiniaceae* umożliwiło mu uzyskanie nowej funkcjonalności.

Interesujący przypadek stanowi białko Hsp201 z bakterii *Edwardisella ictaluri*, należącej do rodziny *Hafniaceae* (Numer akcesji w bazie UniProt C5BHD3). Mimo iż ten ortolog IbpA wchodzi w skład systemu dwubiałkowego, posiada on histydynę na pozycji 66 i kwas glutaminowy na pozycji 109. Sugeruje to, że białko to może potencjalnie stanowić swego rodzaju formę pośrednią między systemem dwubiałkowym i wtórnie pojedynczym IbpA, a jego dokładna charakterystyka biochemiczna mogłaby potencjalnie pomóc lepiej zrozumieć ewolucję IbpA w *Erwiniaceae*.

Substytucje Q66H i G109D, kluczowe dla nabycia nowej funkcjonalności przez wtórnie pojedyncze białko IbpA z *Erwiniaceae*, zlokalizowane są w bezpośredniej bliskości kieszeni β 4- β 8 na domenie ACD (rozdział 7.8.1). Wcześniejsze badania wykazały, że kieszeń ta stanowi miejsce interakcji z regionem CTE (Stróżecka i in. 2012; Reinle, Mogk, i Bukau 2022). W niektórych sHsp stanowi ona, obok regionu NTE, jedno z dwóch głównych miejsc wiążących substrat (Reinle, Mogk, i Bukau 2022).

W moich badaniach pokazałem, że substytucje te wpływają w pewnym stopniu na obie te interakcje – doprowadziły one do obniżenia powinowactwa domeny ACD białka AncA₀ zarówno do badanych substratów, jak i do peptydu C, reprezentującego region CTE tego białka.

W połączeniu z obserwacją, że substytucje te wpływają na zdolność współczesnych białek IbpA z *E. coli* i *E. amylovora* do inhibicji wiązania systemu Hsp70 do zagregowanych substratów opłaszczonych przez sHsp (rozdział 7.9), pozwala to postawić hipotezę, że w toku ewolucji wtórnie pojedynczego białka IbpA w rodzinie *Erwiniaceae* substytucje na pozycjach 66 i 109 doprowadziły do osłabienia oddziaływań stabilizujących zewnętrzną warstwę kompleksu sHsp – substrat. Ułatwiło to systemowi Hsp70 wypieranie sHsp z tego kompleksu w momencie rozpoczęcia procesu dezagregacji. Miałyby one tym samym efekt analogiczny do działania białka IbpB z systemu dwubiałkowego, które ułatwia dysocjację białka IbpA z kompleksu z substratem podczas procesu dezagregacji (Obuchowski *et al.* 2019).

Zgadzałoby się to z wynikami uzyskanymi przez Źwirowskiego i współpracowników, pokazującymi, że region CTE białka IbpA z *E. coli* odgrywa

kluczową rolę w stabilizacji zewnętrznej warstwy kompleksu sHsp – substrat. Wykazali oni, że dodatek wolnego regionu CTE powoduje częściową dysocjację sHsp tworzących tą zewnętrzną warstwę, co umożliwiło zlepianie się kompleksów sHsp – substrat w większe struktury (Żwirowski i in. 2017).

Istotny wpływ siły oddziaływania między domeną ACD, a regionem CTE na stabilizację zewnętrznej warstwy kompleksu sHsp – substrat sugerują również wyniki uzyskane przez Giese i Vierling, dotyczące Hsp 16.6 z cyjanobakterii *Synechocystis sp.* PCC6803. Wprowadzenie do tego białka mutacji wzmacniających oddziaływanie między regionem CTE, a kieszenią β 4- β 8 domeny ACD spowodowało obniżenie jego zdolności do stymulacji reaktywacji lucyferazy przez system Hsp70. Mutacje znacząco osłabiające to oddziaływanie z kolei spowodowały znaczące obniżenie zdolności tego białka do hamowania agregacji (co można interpretować jako upośledzenie sekwestracji agregującego substratu) i całkowitą utratę zdolności do stymulacji reaktywacji lucyferazy, najprawdopodobniej związanej ze zbyt słabą aktywnością sekwestrazy. Wprowadzenie obu typów mutacji, najprawdopodobniej prowadzące do jedynie nieznacznego osłabienia oddziaływania między regionem CTE, a domeną ACD, skutkowało zwiększeniem wydajności stymulacji reaktywacji lucyferazy (Giese i Vierling 2002).

Wyniki te zdają się potwierdzać wnioski postawione w niniejszej pracy sugerujące, że niewielkie osłabienie oddziaływania między regionem CTE i domeną ACD, które zaszło w toku ewolucji wtórnie pojedynczego lbpA z rodziny *Erwiniaceae*, ułatwiło dysocjację sHsp tworzących zewnętrzną warstwę kompleksu sHsp – substrat podczas procesu dezagregacji, zwiększając jego zdolność do efektywnej stymulacji dezagregacji i przywracania konformacji natywnej sekwestrowanego substratu.

Zgadniają się one również z obserwowaną niższą aktywnością sekwestrazy charakteryzującą wtórnie pojedyncze białko lbpA z *E. amylovora* w porównaniu z białkiem lbpA z *E. coli*.

Obserwowany nieznaczny spadek rozmiarów oligomerów tworzonych przez białko AncA₀, jak i ułatwienie jego deoligomeryzacji na skutek wprowadzenia do niego substytucji Q66H i G109D (rozdział 6.8.3) pozostaje w zgodności z kluczową rolą oddziaływania między domeną ACD, a regionem CTE w tworzeniu i stabilizacji oligomerów sHsp (Giese i Vierling 2002; Stróżecka i in. 2012). Niewielki stopień obserwowanego efektu może być tłumaczony względnie niewielkimi zmianami powinowactwa spowodowanymi wprowadzeniem tych substytucji.

Wpływ substytucji na pozycjach 66 i 109 na oddziaływanie między domeną ACD, a substratem również może przyczyniać się do ułatwienia zastępowania sHsp w kompleksie z substratem przez system Hsp70. Oddziaływanie to było osłabione przez substytucje Q66H i G109D zarówno w przypadku, gdy analizowanym substratem była zagregowana lucyferaza, jak i zagregowane białka lizatu *E. coli*. Nie można wykluczyć,

że w pewnym stopniu ułatwiło to dysocjację sHsp związanego z substratem podczas procesu dezagregacji. Należy przy tym pamiętać, że samo wiązanie substratu przez domenę ACD nie jest wystarczające, aby zapewnić aktywność sHsp jako białek opiekuńczych – domeny ACD białek AncA₀, AncA₀ Q66H G109D, jak i współczesnego białka lbpA z *E. coli* pozbawione flankujących regionów CTE i NTE nie były w stanie sekwestrować agregującej lucyferazy ani stymulować jej reaktywacji. Pozostaje to w zgodności z wynikami uzyskanymi przez Stróżecką i wsp., pokazującymi, że delecja zarówno regionu CTE, jak i NTE powoduje utratę zdolności białka lbpA z *E. coli* do sekwestracji agregującej lucyferazy i co za tym idzie, stymulacji jej późniejszej reaktywacji przez system Hsp70 i białka Hsp100 (Stróżecka i in. 2012).

Uzyskane przeze mnie wyniki sugerują, że zdolność do wydajnej, samodzielnej stymulacji dezagregacji i ponownego zwijania sekwestrowanych substratów, która wykształciła się we wtórnie pojedynczym lbpA w *Erwiniaceae*, ma swoją przyczynę w nieznacznym osłabieniu oddziaływań stabilizujących zewnętrzną warstwę kompleksu sHsp – substrat na skutek substytucji, które zaszły w kieszeni β 4- β 8 domeny ACD. Doprowadziło to do ułatwienia wyparcia sHsp z kompleksu z substratem przez system Hsp70 podczas procesu dezagregacji, zmniejszając zapotrzebowanie na współpracę z białkiem lbpB.

Istotne znaczenie może mieć fakt, że efekt ten został uzyskany przez stosunkowo niewielkie osłabienie oddziaływań między domeną ACD, a regionem CTE i substratami. Zbyt silne zaburzenie tych oddziaływań prawdopodobnie mogłoby, jak zaobserwowano np. w pracy Giese i Vierling, doprowadzić do całkowitej utraty funkcjonalności sHsp, które nie byłoby w stanie efektywnie sekwestrować agregujących substratów (Giese i Vierling 2002). Uzyskane wyniki pokazują, że pojedyncze mutacje, powodujące stosunkowo niewielkie zmiany siły oddziaływań międzycząsteczkowych, mogą mieć istotny wpływ na aktywność białek sHsp.

Przeprowadzone badania miały charakter badań podstawowych. Uzyskane przeze mnie wyniki pozwalają lepiej zrozumieć funkcjonowanie bakteryjnych sHsp w procesie dezagregacji białek, zwłaszcza mechanizm ich oddziaływania z substratem. Rzucają też światło na przykład mechanizmu, poprzez który bardziej złożony system składający się z dwóch, współpracujących ze sobą białek może w toku ewolucji ulec uproszczeniu do systemu jednobiałkowego.

9. Bibliografia

- Aevermann, Brian D., i Elizabeth R. Waters. 2008. „A comparative genomic analysis of the small heat shock proteins in *Caenorhabditis elegans* and *briggssae*”. *Genetica* 133 (3): 307–19. <https://doi.org/10.1007/s10709-007-9215-9>.
- Ashkenazy, H., O. Penn, A. Doron-Faigenboim, O. Cohen, G. Cannarozzi, O. Zomer, i T. Pupko. 2012. „FastML: a web server for probabilistic reconstruction of ancestral sequences”. *Nucleic Acids Research* 40 (W1): W580–84. <https://doi.org/10.1093/nar/gks498>.
- Basha, Eman, Kenneth L. Friedrich, i Elizabeth Vierling. 2006. „The N-terminal arm of small heat shock proteins is important for both chaperone activity and substrate specificity”. *Journal of Biological Chemistry* 281 (52): 39943–52. <https://doi.org/10.1074/jbc.M607677200>.
- Benndorf, Rainer, Jody L. Martin, Sergei L. Kosakovsky Pond, i Joel O. Wertheim. 2014. „Neuropathy- and myopathy-associated mutations in human small heat shock proteins: Characteristics and evolutionary history of the mutation sites”. *Mutation Research - Reviews in Mutation Research* 761: 15–30. <https://doi.org/10.1016/j.mrrev.2014.02.004>.
- Bepperling, Alexander, Ferdinand Alte, Thomas Kriehuber, Nathalie Braun, Sevil Weinkauff, Michael Groll, Martin Haslbeck, i Buchner Johannes. 2012. „Alternative bacterial two-component small heat shock protein systems”. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.1209565109>.
- Bertelsena, Eric B., Lyra Chang, Jason E. Gestwicki, i Erik R.P. Zuiderweg. 2009. „Solution conformation of wild-type *E. coli* Hsp70 (DnaK) chaperone complexed with ADP and substrate”. *Proceedings of the National Academy of Sciences of the United States of America* 106 (21): 8471–76. <https://doi.org/10.1073/pnas.0903503106>.
- Bourrelle-Langlois, Maxime, Geneviève Morrow, Stéphanie Finet, i Robert M. Tanguay. 2016. „In vitro structural and functional characterization of the small heat shock proteins (sHSP) of the cyanophage S-ShM2 and its host, *synechococcus* sp. WH7803”. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0162233>.
- Chamera, Tomasz, Agnieszka Kłosowska, Anna Janta, Hubert Wyszowski, Igor Obuchowski, Krzysztof Gumowski, i Krzysztof Liberek. 2019. „Selective Hsp70-Dependent Docking of Hsp104 to Protein Aggregates Protects the Cell from the Toxicity of the Disaggregase”. *Journal of Molecular Biology* 431 (11): 2180–96. <https://doi.org/10.1016/j.jmb.2019.04.014>.
- Clark, A. R., C. E. Naylor, C. Bagnéris, N. H. Keep, i C. Slingsby. 2011. „Crystal structure of R120G disease mutant of human α -crystallin domain dimer shows closure of a groove”. *Journal of Molecular Biology* 408 (1): 118–34. <https://doi.org/10.1016/j.jmb.2011.02.020>.
- Cohen, Ofir, i Tal Pupko. 2011. „Inference of Gain and Loss Events from Phyletic Patterns Using Stochastic Mapping and Maximum Parsimony—A Simulation Study”. *Genome Biology and Evolution* 3 (styczeń): 1265–75. <https://doi.org/10.1093/gbe/evr101>.
- Cohen, Ofir, Nimrod D Rubinstein, Adi Stern, Uri Gophna, i Tal Pupko. 2008. „A likelihood framework to analyse phyletic patterns”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1512): 3903–11. <https://doi.org/10.1098/rstb.2008.0177>.
- De Jong, W. W., J. A.M. Leunissen, i C. E.M. Voorter. 1993. „Evolution of the α -crystallin/small heat-shock protein family”. *Molecular Biology and Evolution* 10 (1): 103–26. <https://doi.org/10.1093/oxfordjournals.molbev.a039992>.

- Dobson, Christopher M. 2004. „Principles of protein folding, misfolding and aggregation”. *Seminars in Cell and Developmental Biology* 15 (1): 3–16. <https://doi.org/10.1016/j.semcdb.2003.12.008>.
- Dutkiewicz, Rafal, i Malgorzata Nowak. 2018. „Molecular chaperones involved in mitochondrial iron–sulfur protein biogenesis”. *Journal of Biological Inorganic Chemistry* 23 (4): 569–79. <https://doi.org/10.1007/s00775-017-1504-x>.
- Eick, Geeta N., Jamie T. Bridgham, Douglas P. Anderson, Michael J. Harms, i Joseph W. Thornton. 2016. „Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty”. *Molecular Biology and Evolution* 34 (2): msw223. <https://doi.org/10.1093/molbev/msw223>.
- Ellis, R. John. 1997. „Molecular chaperones: Avoiding the crowd”. *Current Biology* 7 (9): R531–33. [https://doi.org/10.1016/S0960-9822\(06\)00273-9](https://doi.org/10.1016/S0960-9822(06)00273-9).
- Ellis, R. John, i Allen P. Minton. 2006. „Protein Aggregation in Crowded Environments” 387 (5): 485–97. <https://doi.org/10.1515/BC.2006.064>.
- Ernst, Katharina, Johannes Schmid, Matthias Beck, Marlen Hägele, Meike Hohwieler, Patricia Hauff, Anna Katharina Ückert, i in. 2017. „Hsp70 facilitates trans-membrane transport of bacterial ADP-ribosylating toxins into the cytosol of mammalian cells”. *Scientific Reports* 7 (1): 1–16. <https://doi.org/10.1038/s41598-017-02882-y>.
- Evans, Richard, Michael O’Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Židek, i in. 2022. „Protein Complex Prediction with AlphaFold-Multimer”. bioRxiv. <https://doi.org/10.1101/2021.10.04.463034>.
- Fuchs, Margit, Dominic J. Poirier, Samuel J. Seguin, Herman Lambert, Serena Carra, Steve J. Charette, i Jacques Landry. 2009. „Identification of the key structural motifs involved in HspB8/HspB6–Bag3 interaction”. *Biochemical Journal* 425 (1): 245–57. <https://doi.org/10.1042/BJ20090907>.
- Giese, Kim C, i Elizabeth Vierling. 2002. „Changes in Oligomerization Are Essential for the Chaperone Activity of a Small Heat Shock Protein in Vivo and in Vitro *”. *Journal of Biological Chemistry* 277 (48): 46310–18. <https://doi.org/10.1074/jbc.M208926200>.
- Goloubinoff, Pierre, i Paolo De Los Rios. 2007. „The mechanism of Hsp70 chaperones: (entropic) pulling the models together”. *Trends in Biochemical Sciences* 32 (8): 372–80. <https://doi.org/10.1016/j.tibs.2007.06.008>.
- Gonçalves, Conrado C., Itai Sharon, T. Martin Schmeing, Carlos H.I. Ramos, i Jason C. Young. 2021. „The chaperone HSPB1 prepares protein aggregates for resolubilization by HSP70”. *Scientific Reports* 11 (1): 1–16. <https://doi.org/10.1038/s41598-021-96518-x>.
- Harms, Michael J., i Joseph W. Thornton. 2010. „Analyzing protein structure and function using ancestral gene reconstruction”. *Current Opinion in Structural Biology* 20 (3): 360–66. <https://doi.org/10.1016/j.sbi.2010.03.005>.
- Haslbeck, Martin, Nathalie Braun, Thusnelda Stromer, Bettina Richter, Natascha Model, Sevil Weinkauff, i Johannes Buchner. 2004. „Hsp42 is the general small heat shock protein in the cytosol of *Saccharomyces cerevisiae*”. *EMBO Journal* 23 (3): 638–49. <https://doi.org/10.1038/sj.emboj.7600080>.
- Haslbeck, Martin, i Elizabeth Vierling. 2015. „A first line of stress defense: Small heat shock proteins and their function in protein homeostasis”. *Journal of Molecular Biology* 427 (7): 1537–48. <https://doi.org/10.1016/j.jmb.2015.02.002>.
- Haslbeck, Martin, Sevil Weinkauff, i Johannes Buchner. 2019. „Small heat shock proteins: Simplicity meets complexity”. *Journal of Biological Chemistry* 294 (6): 2121–32. <https://doi.org/10.1074/jbc.REV118.002809>.
- Hilario, Eduardo, Francisco Javier Medrano Martin, Maria Célia Bertolini, i Li Fan. 2011. „Crystal structures of xanthomonas small heat shock protein provide a structural basis for an active molecular chaperone oligomer”. *Journal of Molecular Biology* 408 (1): 74–86. <https://doi.org/10.1016/j.jmb.2011.02.004>.

- Hu, Yanmin, Farahnaz Movahedzadeh, Neil G. Stoker, i Anthony R.M. Coates. 2006. „Deletion of the Mycobacterium tuberculosis α -crystallin-like hspX gene causes increased bacterial growth in vivo”. *Infection and Immunity* 74 (2): 861–68. <https://doi.org/10.1128/IAI.74.2.861-868.2006>.
- Huang, Jing, Sarah Rauscher, Grzegorz Nawrocki, Ting Ran, Michael Feig, Bert L. de Groot, Helmut Grubmüller, i Alexander D. MacKerell. 2017. „CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins”. *Nature methods* 14 (1): 71–73. <https://doi.org/10.1038/nmeth.4067>.
- Ingolia, T. D., i E. A. Craig. 1982. „Four small Drosophila heat shock proteins are related to each other and mammalian α -crystallin”. *Proceedings of the National Academy of Sciences of the United States of America* 79 (7 1): 2360–64. <https://doi.org/10.1073/pnas.79.7.2360>.
- Jahn, Thomas R., i Sheena E. Radford. 2008. „Folding versus aggregation: Polypeptide conformations on competing pathways”. *Archives of Biochemistry and Biophysics* 469 (1): 100–117. <https://doi.org/10.1016/j.abb.2007.05.015>.
- Jaya, Nomalie, Victor Garcia, i Elizabeth Vierling. 2009. „Substrate binding site flexibility of the small heat shock protein molecular chaperones”. *Proceedings of the National Academy of Sciences of the United States of America* 106 (37): 15604–9. <https://doi.org/10.1073/pnas.0902177106>.
- Jeffares, Daniel C, Bartłomiej Tomiczek, Victor Sojo, i Mario Reis. 2015. „A worked example of estimating omega and testing for adaptive evolution in six parasite species”. *Methods Mol Biol.* 1201: 65–69.
- Jones, David T., William R. Taylor, i Janet M. Thornton. 1992. „The rapid generation of mutation data matrices from protein sequences”. *Bioinformatics* 8 (3): 275–82. <https://doi.org/10.1093/bioinformatics/8.3.275>.
- Kalyaanamoorthy, Subha, Bui Quang Minh, Thomas K F Wong, Arndt von Haeseler, i Lars S Jermiin. 2017. „ModelFinder: fast model selection for accurate phylogenetic estimates”. *Nature Methods* 14 (6): 587–89. <https://doi.org/10.1038/nmeth.4285>.
- Kennaway, Christopher K., Justin L.P. Benesch, Ulrich Gohlke, Luchun Wang, Carol V. Robinson, Elena V. Orlova, Helen R. Saibi, i Nicholas H. Keep. 2005. „Dodecameric structure of the small heat shock protein Acr1 from Mycobacterium tuberculosis”. *Journal of Biological Chemistry* 280 (39): 33419–25. <https://doi.org/10.1074/jbc.M504263200>.
- Kriehuber, Thomas, Thomas Rattei, Thomas Weinmaier, Alexander Bepperling, Martin Haslbeck, i Johannes Buchner. 2010. „Independent evolution of the core domain and its flanking sequences in small heat shock proteins”. *FASEB Journal* 24 (10): 3633–42. <https://doi.org/10.1096/fj.10-156992>.
- Krukenberg, Kristin A., Friedrich Förster, Luke M. Rice, Andrej Sali, i David A. Agard. 2008. „Multiple Conformations of E. coli Hsp90 in Solution: Insights into the Conformational Dynamics of Hsp90”. *Structure* 16 (5): 755–65. <https://doi.org/10.1016/j.str.2008.01.021>.
- Kuczyńska-Wiśnik, Dorota, Sabina Kędzierska, Ewelina Matuszewska, Peter Lund, Alina Taylor, Barbara Lipińska, i Ewa Laskowska. 2002. „The Escherichia coli small heat-shock proteins lbpA and lbpB prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock”. *Microbiology* 148 (6): 1757–65. <https://doi.org/10.1099/00221287-148-6-1757>.
- Lee, Garrett J., Alan M. Roseman, Helen R. Saibil, i Elizabeth Vierling. 1997. „A small heat shock protein stably binds heat-denatured model substrates and can maintain a substrate in a folding-competent state”. *The EMBO Journal* 16 (3): 659–71. <https://doi.org/10.1093/emboj/16.3.659>.

- Liberek, Krzysztof, Agnieszka Lewandowska, i Szymon Ziętkiewicz. 2008. „Chaperones in control of protein disaggregation”. *EMBO Journal* 27 (2): 328–35. <https://doi.org/10.1038/sj.emboj.7601970>.
- Lim, Joo Hyun, Falk Martin, Bernard Guiard, Nikolaus Pfanner, i Wolfgang Voos. 2001. „The mitochondrial Hsp70-dependent import system actively unfolds preproteins and shortens the lag phase of translocation”. *EMBO Journal* 20 (5): 941–50. <https://doi.org/10.1093/emboj/20.5.941>.
- Lipińska, Natalia, Szymon Ziętkiewicz, Alicja Sobczak, Agnieszka Jurczyk, Wojciech Potocki, Ewa Morawiec, Aleksandra Wawrzycka, i in. 2013. „Disruption of ionic interactions between the nucleotide binding domain 1 (NBD1) and middle (M) domain in Hsp100 disaggregase unleashes toxic hyperactivity and partial independence from Hsp70”. *Journal of Biological Chemistry* 288 (4): 2857–69. <https://doi.org/10.1074/jbc.M112.387589>.
- Mani, Nandini, Spraha Bhandari, Rodolfo Moreno, Liya Hu, B. V. Venkataram Prasad, i Kaza Suguna. 2016. „Multiple oligomeric structures of a bacterial small heat shock protein”. *Scientific Reports* 6 (March): 1–12. <https://doi.org/10.1038/srep24019>.
- Mayer, M. P., i B. Bukau. 2005. „Hsp70 chaperones: Cellular functions and molecular mechanism”. *Cellular and Molecular Life Sciences* 62 (6): 670–84. <https://doi.org/10.1007/s00018-004-4464-6>.
- McClellan, Amie J., Yu Xia, Adam M. Deutschbauer, Ron W. Davis, Mark Gerstein, i Judith Frydman. 2007. „Diverse Cellular Functions of the Hsp90 Molecular Chaperone Uncovered Using Systems Approaches”. *Cell* 131 (1): 121–35. <https://doi.org/10.1016/j.cell.2007.07.036>.
- Meyer, Anne S., Joel R. Gillespie, Dirk Walther, Ian S. Millet, Sebastian Doniach, i Judith Frydman. 2003. „Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis”. *Cell* 113 (3): 369–81. [https://doi.org/10.1016/S0092-8674\(03\)00307-6](https://doi.org/10.1016/S0092-8674(03)00307-6).
- Mirdita, Milot, Konstantin Schütze, Yoshitaka Moriwaki, Lim Heo, Sergey Ovchinnikov, i Martin Steinegger. 2022. „ColabFold: Making Protein Folding Accessible to All”. *Nature Methods* 19 (6): 679–82. <https://doi.org/10.1038/s41592-022-01488-1>.
- Mogk, Axel, Elke Deuerling, Sonja Vorderwülbecke, Elizabeth Vierling, i Bernd Bukau. 2003. „Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation”. *Molecular Microbiology* 50 (2): 585–95. <https://doi.org/10.1046/j.1365-2958.2003.03710.x>.
- Mogk, Axel, Eva Kummer, i Bernd Bukau. 2015. „Cooperation of Hsp70 and Hsp100 chaperone machines in protein disaggregation”. *Frontiers in Molecular Biosciences* 2 (MAY): 1–10. <https://doi.org/10.3389/fmolb.2015.00022>.
- Moreau, Kate L, i Jonathan A King. 2012. „Protein misfolding and aggregation in cataract disease and prospects for prevention”. *Trends in Molecular Medicine* 18 (5): 273–82. <https://doi.org/10.1016/j.molmed.2012.03.005>.
- Münchbach, Martin, Andreas Nocker, i Franz Narberhaus. 1999. „Multiple small heat shock proteins in rhizobia”. *Journal of Bacteriology* 181 (1): 83–90. <https://doi.org/10.1128/jb.181.1.83-90.1999>.
- Natarajan, Chandrasekhar, Anthony V. Signore, Naim M. Bautista, Federico G. Hoffmann, Jeremy R. H. Tame, Angela Fago, i Jay F. Storz. 2023. „Evolution and Molecular Basis of a Novel Allosteric Property of Crocodilian Hemoglobin”. *Current Biology* 33 (1): 98-108.e4. <https://doi.org/10.1016/j.cub.2022.11.049>.
- Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, i Bui Quang Minh. 2015. „IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies”. *Molecular Biology and Evolution* 32 (1): 268–74. <https://doi.org/10.1093/molbev/msu300>.

- Obuchowski, Igor, Piotr Karaś, i Krzysztof Liberek. 2021. „The Small Ones Matter—sHsps in the Bacterial Chaperone Network”. *Frontiers in Molecular Biosciences* 8 (May): 1–7. <https://doi.org/10.3389/fmolb.2021.666893>.
- Obuchowski, Igor, Artur Piróg, Milena Stolarska, Bartłomiej Tomiczek, i Krzysztof Liberek. 2019. „Duplicate divergence of two bacterial small heat shock proteins reduces the demand for Hsp70 in refolding of substrates”. Zredagowane przez Danielle A. Garsin. *PLoS Genetics* 15 (10): e1008479. <https://doi.org/10.1371/journal.pgen.1008479>.
- Palmer, Marike, Emma T. Steenkamp, Martin P.A. Coetzee, Juanita R. Avontuur, Wai Yin Chan, Elritha van Zyl, Jochen Blom, i Stephanus N. Venter. 2018. „Mixta gen. Nov., a new genus in the Erwiniaceae”. *International Journal of Systematic and Evolutionary Microbiology* 68 (4): 1396–1407. <https://doi.org/10.1099/ijsem.0.002540>.
- Pearl, Laurence H. 2016. „Review: The HSP90 molecular chaperone - An enigmatic ATPase”. *Biopolymers* 105 (8): 594–607. <https://doi.org/10.1002/bip.22835>.
- Piróg, Artur, Francesca Cantini, Łukasz Nierzwicki, Igor Obuchowski, Bartłomiej Tomiczek, Jacek Czub, i Krzysztof Liberek. 2021. „Two bacterial small heat shock proteins, IbpA and IbpB, form a functional heterodimer”. *Journal of Molecular Biology*, 167054. <https://doi.org/10.1016/j.jmb.2021.167054>.
- Pupko, T., I. Pe'er, M. Hasegawa, D. Graur, i N. Friedman. 2002. „A branch-and-bound algorithm for the inference of ancestral amino-acid sequences when the replacement rate varies among sites: Application to the evolution of five gene families”. *Bioinformatics* 18 (8): 1116–23. <https://doi.org/10.1093/bioinformatics/18.8.1116>.
- Ratajczak, Elzbieta, Joanna Stróżecka, Marlena Matuszewska, Szymon Zietkiewicz, Dorota Kuczyńska-Wiśnik, Ewa Laskowska, i Krzysztof Liberek. 2010. „IbpA the small heat shock protein from *Escherichia coli* forms fibrils in the absence of its cochaperone IbpB”. *FEBS Letters* 584 (11): 2253–57. <https://doi.org/10.1016/j.febslet.2010.04.060>.
- Ratajczak, Elzbieta, Szymon Zietkiewicz, i Krzysztof Liberek. 2009. „Distinct Activities of *Escherichia coli* Small Heat Shock Proteins IbpA and IbpB Promote Efficient Protein Disaggregation”. *Journal of Molecular Biology* 386 (1): 178–89. <https://doi.org/10.1016/j.jmb.2008.12.009>.
- Reinle, Kevin, Axel Mogk, i Bernd Bukau. 2022. „The Diverse Functions of Small Heat Shock Proteins in the Proteostasis Network: Functions and mechanisms of sHsps”. *Journal of Molecular Biology* 434 (1): 167157. <https://doi.org/10.1016/j.jmb.2021.167157>.
- Siddique, Masood, Sascha Gernhard, Pascal von Koskull-Döring, Elizabeth Vierling, i Klaus-Dieter Scharf. 2008. „The plant sHSP superfamily: five new members in *Arabidopsis thaliana* with unexpected properties”. *Cell Stress and Chaperones* 13 (2): 183–97. <https://doi.org/10.1007/s12192-008-0032-6>.
- Sievers, Fabian, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, i in. 2011. „Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega”. *Molecular Systems Biology* 7 (październik): 539. <https://doi.org/10.1038/msb.2011.75>.
- Simmons, Mark P, i Helga Ochoterena. 2000. „Gaps as Characters in Sequence-Based Phylogenetic Analyses”. *Systematic Biology* 49 (2): 369–81.
- Stengel, Florian, Andrew J. Baldwin, Alexander J. Painter, Nomalie Jaya, Eman Basha, Lewis E. Kay, Elizabeth Vierling, Carol V. Robinson, i Justin L.P. Benesch. 2010. „Quaternary dynamics and plasticity underlie small heat shock protein chaperone function”. *Proceedings of the National Academy of Sciences of the United States of America* 107 (5): 2007–12. <https://doi.org/10.1073/pnas.0910126107>.

- Stróżecka, Joanna, Elzbieta Chruściel, Emilia Górna, Aneta Szymańska, Szymon Ziętkiewicz, i Krzysztof Liberek. 2012. „Importance of N- and C-terminal regions of IbpA, Escherichia coli small heat shock protein, for chaperone function and oligomerization”. *Journal of Biological Chemistry* 287 (4): 2843–53. <https://doi.org/10.1074/jbc.M111.273847>.
- Sun, Yu, i Thomas H. MacRae. 2005. „The small heat shock proteins and their role in human disease”. *FEBS Journal* 272 (11): 2613–27. <https://doi.org/10.1111/j.1742-4658.2005.04708.x>.
- Tang, Yun Chi, Hung Chun Chang, Annette Roeben, Dirk Wischnewski, Nadine Wischnewski, Michael J. Kerner, F. Ulrich Hartl, i Manajit Hayer-Hartl. 2006. „Structural Features of the GroEL-GroES Nano-Cage Required for Rapid Folding of Encapsulated Protein”. *Cell* 125 (5): 903–14. <https://doi.org/10.1016/j.cell.2006.04.027>.
- Tomoyasu, Toshifumi, Atsushi Tabata, Yoko Ishikawa, Robert A. Whitley, i Hideaki Nagamune. 2013. „Small heat shock protein AgsA: An effective stabilizer of enzyme activities”. *Journal of Bioscience and Bioengineering* 115 (1): 15–19. <https://doi.org/10.1016/j.jbiosc.2012.08.001>.
- Ungelenk, Sophia, Fatemeh Moayed, Chi Ting Ho, Tomas Grousl, Annette Scharf, Alireza Mashaghi, Sander Tans, Matthias P. Mayer, Axel Mogk, i Bernd Bukau. 2016. „Small heat shock proteins sequester misfolding proteins in near-native conformation for cellular protection and efficient refolding”. *Nature Communications* 7: 1–14. <https://doi.org/10.1038/ncomms13673>.
- Vabulas, R. Martin, Swasti Raychaudhuri, Manajit Hayer-Hartl, i F. Ulrich Hartl. 2010. „Protein Folding in the Cytoplasm and the Heat Shock Response”. *Cold Spring Harbor Perspectives in Biology* 2 (12): a004390–a004390. <https://doi.org/10.1101/cshperspect.a004390>.
- Van Der Spoel, David, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, i Herman J. C. Berendsen. 2005. „GROMACS: Fast, Flexible, and Free”. *Journal of Computational Chemistry* 26 (16): 1701–18. <https://doi.org/10.1002/jcc.20291>.
- Walker, Glenda A., i Gordon J. Lithgow. 2003. „Lifespan extension in *C. elegans* by a molecular chaperone dependent upon insulin-like signals.” *Aging cell* 2 (2): 131–39. <https://doi.org/10.1046/j.1474-9728.2003.00045.x>.
- Weibezahn, Jimena, Peter Tessarz, Christian Schlieker, Regina Zahn, Zeljka Maglica, Sukyeong Lee, Hanswalter Zentgraf, i in. 2005. „Thermotolerance Requires Refolding of Aggregated Proteins by Substrate Translocation through the Central Pore of ClpB Jimena”. *Trends in Biochemical Sciences* 30 (3): i. [https://doi.org/10.1016/s0968-0004\(05\)00043-5](https://doi.org/10.1016/s0968-0004(05)00043-5).
- Yang, Ziheng. 1998. „Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution”. *Molecular Biology and Evolution* 15 (5): 568–73. <https://doi.org/10.1093/oxfordjournals.molbev.a025957>.
- . 2007. „PAML 4: Phylogenetic analysis by maximum likelihood”. *Molecular Biology and Evolution* 24 (8): 1586–91. <https://doi.org/10.1093/molbev/msm088>.
- Yang, Ziheng, i Rasmus Nielsen. 2002. „Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages”. *Molecular Biology and Evolution* 19 (6): 908–17. <https://doi.org/10.1093/oxfordjournals.molbev.a004148>.
- Yuan, Ying, Deborah D. Crane, i Clifton E. Barry. 1996. „Stationary phase-associated protein expression in Mycobacterium tuberculosis: Function of the mycobacterial α -crystallin homolog”. *Journal of Bacteriology* 178 (15): 4484–92. <https://doi.org/10.1182/blood-2011-06-360255> [pii].
- Zolkiewski, Michal, Ting Zhang, i Maria Nagy. 2012. „Aggregate reactivation mediated by the Hsp100 chaperones”. *Archives of Biochemistry and Biophysics* 520 (1): 1–6. <https://doi.org/10.1016/j.abb.2012.01.012>.

Żwirowski, Szymon, Agnieszka Kłosowska, Igor Obuchowski, Nadinath B Nillegoda, Artur Piróg, Szymon Ziętkiewicz, Bernd Bukau, Axel Mogk, i Krzysztof Liberek. 2017. „Hsp70 displaces small heat shock proteins from aggregates to initiate protein refolding”. *The EMBO Journal*. <https://doi.org/10.15252/emj.201593378>.