

Międzyuczelniany Wydział Biotechnologii
Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego

mgr Inga Blanka Jamrozek

Opracowanie modelu *in silico* receptora
kiełkowania przetrwalników GerA
Bacillus subtilis

Developing an *in silico* model of the GerA *Bacillus subtilis*
germination receptor

Praca przedstawiona
Radzie Dyscypliny Nauki Biologiczne Uniwersytetu Gdańskiego
celem uzyskania stopnia doktora w dziedzinie nauk ścisłych i przyrodniczych
w dyscyplinie nauki biologiczne

Promotor: prof. dr hab. Michał Obuchowski
Promotor pomocniczy: dr hab. Rajmund Kaźmierkiewicz, prof. UG

Pracownia Symulacji Układów Biomolekularnych

Gdańsk 2023

*Dziękuję Marcie i Małgosi za wspólną walkę,
podnoszenie się na duchu
i ogrom wsparcia.*

*Dziękuję wszystkim,
których zaangażowanie przyczyniło się do powstania tej pracy:
Promotorom — za poświęcony czas;
Annie Grela i Marcie Hubisz — za wspólną pracę i nieoceniony wkład w badania;
Adamowi Iwanickiemu — za trafne uwagi.*

*Dziękuję też tym, którzy wspierali mnie dobrym słowem
i „trzymali kciuki” przez te długie lata.*

Tacie — za wsparcie i entuzjazm.

Sobie — za wszystko.

Oświadczenie

Ja, niżej podpisana, oświadczam, że przedłożona praca dyplomowa została wykonana przeze mnie samodzielnie, nie narusza praw autorskich, interesów prawnych i materialnych innych osób.

Data

Podpis

Spis treści

1. STRESZCZENIE	7
2. WYKAZ SKRÓTÓW	9
3. WSTĘP TEORETYCZNY.....	10
3.1. Struktura białek	10
3.1.1. Rola struktur i ich znaczenie.....	11
3.1.2. Teoretyczne metody przewidywania struktur białek.....	12
3.2. Modelowanie homologiczne	13
3.2.1. Modelowanie na wielu szablonach	14
3.2.2. RMSD jako wyznacznik podobieństwa strukturalnego.....	15
3.2.3. Strategie oceny jakości modelu	16
3.3. Oddziaływania białek i dokowanie	16
3.3.1. Dokowanie białko-ligand	17
3.3.2. Dokowanie białko-białko.....	19
3.3.3. Obliczeniowe metody przewidywania regionów wiązania	20
3.4. Symbioza metod obliczeniowych i eksperymentalnych.....	21
3.5. <i>Bacillus subtilis</i>	21
3.5.1. Budowa i właściwości przetrwalników	22
3.6. Kielkowanie przetrwalników.....	23
3.7. Receptory kielkowania	24
3.7.1. Receptor kielkowania GerA.....	25
3.7.2. Interakcje receptora z germinantami	26
3.8. Germinosom	27
3.8.1. Rola białka GerD w procesie kielkowania.....	27
3.8.2. Organizacja receptorów w procesie chemotaksji	28
3.9. Błony lipidowe	29
3.9.1. Modelowanie błon fosfolipidowych	30
3.9.2. Model błon biologicznych <i>B. subtilis</i>	31
3.10. Znaczenie badań nad strukturą receptora GerA	33
4. CEL PRACY.....	33
5. MATERIAŁY I METODY	34
5.1. Przewidywanie struktur drugorzędowych.....	34
5.2. Przewidywanie regionów transbłonowych	34
5.3. Wyszukiwanie homologów i zestawiania sekwencji.....	34
5.4. Przewidywanie struktur trzeciorzędowych	35
5.5. Ocena jakości modeli teoretycznych	36

5.6.	Wizualizacja modeli	37
5.7.	Dokowania molekularne.....	37
5.7.1.	Wstępne analizy	37
5.7.2.	Dokowanie ligand-receptor	37
5.7.3.	Dokowania białko-białko	40
5.7.4.	Wizualizacja oddziaływań	41
5.8.	Budowanie modelu błony fosfolipidowej.....	42
5.8.1.	Badania eksperymentalne	42
5.8.2.	Orientacja białek receptora w błonie	42
5.8.3.	Charmm-GUI	43
6.	WYNIKI	45
6.1.	Modele podjednostek receptora GerA.....	45
6.1.1.	Podjednostka A	45
6.1.2.	Model podjednostki B	55
6.1.3.	Model podjednostki C	60
6.1.4.	Podsumowanie — modele podjednostek receptora GerA.....	65
6.2.	Model białka GerD.....	65
6.2.1.	Monomer GerD.....	66
6.2.2.	Trimer GerD.....	68
6.3.	Mutacje w białku GerAA.....	69
6.3.1.	Pozycje 299 i 302	70
6.3.2.	Arginina 347	72
6.3.3.	Motyw PFPP.....	74
6.4.	Modele receptora.....	78
6.4.1.	Wyniki przewidywania miejsc interakcji białko-białko	79
6.4.2.	Dimery podjednostek.....	79
6.4.3.	Kompleks monomerów białek błonowych receptora GerA	96
6.4.4.	Trimery podjednostek	101
6.5.	Dokowanie germinantów.....	103
6.5.1.	Dokowania germinantów do podjednostki A	103
6.5.2.	Dokowania germinantów do podjednostki B	111
6.5.3.	Oddziaływania germinantów z kompleksem receptora GerA zbudowanym z monomerów.....	116
6.5.4.	Oddziaływania germinantów z kompleksem receptora GerA zbudowanym z dimerów 120	
6.5.5.	Podsumowanie wyników dokowań globalnych.....	123
6.6.	Orientacja białek receptora GerA w błonie	125
6.6.1.	Orientacja monomerów w błonie.....	126

6.6.2.	Orientacja dimerów w błonie	127
6.6.3.	Orientacja kompleksów w błonie	128
6.6.4.	Model wewnętrznej błony przetrwalnika	130
7.	DYSKUSJA	133
7.1.	Model receptora GerA	135
7.2.	Wiązanie germinantów	136
8.	BIBLIOGRAFIA	138
9.	ZAŁĄCZNIKI	144
	Załącznik nr 1	144
	Załącznik nr 2	145
	Załącznik nr 3	146

1. STRESZCZENIE

Przetrwalniki bakterii *Bacillus subtilis* są nieaktywnymi metabolicznie formami spoczynkowymi i wyróżniają się wysoką odpornością na szkodliwe czynniki środowiskowe. Gdy warunki środowiskowe stają się bardziej korzystne inicjowany jest proces kiełkowania przetrwalnika, który prowadzi do powstania formy wegetatywnej. Sygnałem biochemicznym dla rozpoczęcia germinacji może być wzrost stężenia specyficznych induktorów — germinantów — w środowisku spora. Inicjacja procesu przebiega przy udziale receptorów kiełkowania.

Jednym ze scharakteryzowanych receptorów kiełkowania jest receptor GerA, który był głównym przedmiotem niniejszej pracy badawczej. Receptor GerA indukuje proces kiełkowania w odpowiedzi na L-alaninę i/lub L-walinę i zbudowany jest z trzech podjednostek: A, B i C. Dane literaturowe pokazują, że wszystkie podjednostki są niezbędne do prawidłowego funkcjonowania receptora. Receptor GerA, na którym skupia się ta praca doktorska zlokalizowany jest w wewnętrznej błonie przetrwalnika, a dwie z jego podjednostek – GerAA i GerAB – są integralnymi białkami błonowymi. Struktura receptora oraz mechanizm przekazywania sygnału inicjującego proces germinacji, pomimo licznych badań, jest nieznaną.

Głównym celem pracy było opracowanie modelu teoretycznego receptora GerA oraz zbadanie potencjalnych miejsc oddziaływania germinantów inicjujących proces kiełkowania. Model ten, biorąc pod uwagę silny związek między strukturą a funkcją biologiczną makromolekuł, może posłużyć do dalszych badań, mających na celu wyjaśnienie mechanizmu inicjacji kiełkowania.

W pierwszym kroku zbudowałam modele teoretyczne podjednostek receptora (białek GerAA, GerAB oraz GerAC). Przeprowadziłam modelowanie homologiczne cząsteczek białka. Uzyskane modele podjednostek kompleksu posłużyły mi następnie do zbudowania modeli receptora oraz określenia potencjalnych miejsc oddziaływania germinantów. Przygotowałam również modele mutantów białka GerAA, za pomocą których wyjaśniłam wpływ struktury na fenotyp mutantów.

Aby przewidzieć potencjalne miejsca oddziaływania germinantów z receptorem wykorzystałam metody dokowania molekularnego. Ponieważ w literaturze nie ma wzmianek o tym, gdzie dokładnie znajduje się kieszeń wiążąca ligandy, a jedynie mówi się ogólnie o podjednostkach A i B jako potencjalnie wiążących germinanty, w badaniach wykorzystałam dokowanie globalne, które dostarczyło informacji o lokalizacji potencjalnych miejsc oddziaływania.

Na podstawie obserwacji podobieństw procesu łączenia się receptorów chemotaksji oraz receptorów kiełkowania, a także homologów użytych do modelowania podjednostek, podczas badań przyjął, że podjednostki mogą formować homodimery lub homotrimery, które następnie, przy udziale białka GerD łączą się, tworząc w receptor.

Na podstawie uzyskanych wyników zaprezentowane zostały dwa modele receptora GerA. Pierwszy zbudowany z wykorzystaniem monomerów podjednostek, natomiast drugi model zbudowany jest z dimerów, a zatem tworzy, podobnie jak receptory chemotaksji, trimer dimerów. Dokowanie germinantów do podjednostek oraz kompleksu receptora wskazuje kilka potencjalnych miejsc interakcji, które odpowiadają za inicjację procesu kiełkowania.

Bacillus subtilis spores are metabolically inactive and are distinguished by their high resistance to harmful environmental factors. When environmental conditions become more favorable, the spore germination process is initiated, leading to the formation of a vegetative form. The biochemical signal for the initiation of germination may be an increase in the concentration of specific inducers - germinants - in the spore environment. The initiation of the process takes place with the participation of germination receptors.

One of the germination receptors is the GerA receptor, which was the main focus of this research work. The GerA receptor induces the germination process in response to L-alanine and/or L-valine and is built of three subunits: A, B and C. Literature data show that all subunits are essential for the receptor to function properly. The GerA receptor, which is the focus of this dissertation, is located in the inner membrane of the spore and two of its subunits, GerAA and GerAB, are integral membrane proteins. The structure of the receptor and the mechanism of signal transduction initiating the germination process, despite numerous studies, remains unknown.

The main aim of this study was to develop a theoretical model of the GerA receptor and to investigate potential binding sites for germinants initiating the germination process. This model, given the strong relationship between structure and biological function of macromolecules, can be used for further studies aimed at elucidating the mechanism of germination initiation.

In a first step, theoretical models of the receptor subunits (GerAA, GerAB and GerAC proteins) were built. Modelling was performed using homology modelling methods. The obtained models were then used to build receptor models and to determine potential binding sites of germinants. Mutant models of GerAA protein were also investigated and then were used to explain the influence of structure on the phenotype of the mutants.

To predict potential binding sites of germinants, molecular docking methods were used. Since the literature does not mention where exactly the ligand-binding pocket is located and only speaks in general terms of the A and B subunits as potentially binding germinants, global docking was used in this study to provide information on the location of potential binding sites.

Based on the observation of similarities in the clustering process of chemotaxis and germination receptors, as well as the homologues used to model the subunits, it was assumed during the study that the subunits could form homodimers or homotrimers, which then, with the involvement of the GerD protein, cluster into a receptor.

Based on the results obtained during research, two models of the GerA receptor were presented. The first one was built on the basis of monomers of subunits, whereas the second model is built of dimers and thus forms, similarly to chemotaxis receptors, a trimer of dimers. Docking of germinants to the subunits and the receptor complex indicates several potential interaction sites that are responsible for the initiation of the germination process.

2. WYKAZ SKRÓTÓW

aa — reszta aminokwasowa

RMSD — ang. *root-mean-square deviation* (odchylenie średniokwadratowe)

CASP — Critical Assessment of protein Structure Prediction

PDB — Protein Data Bank

DM — Dynamika Molekularna

PSA — ang. Pairwise Sequence Alignment

MSA — ang. Multiple Sequence Alignment

DPA — ang. Dipicolinic acid (Kwas dipikolinowy)

NMR — ang. *Nuclear Magnetic Resonance* (spektroskopia magnetycznego rezonansu jądrowego)

GDT — ang. *Global Distance Test* (wartość odległości globalnych)

MQAP — ang. *Model Quality Assessment Program*

PPI — ang. *Protein-Protein Interaction* (interakcje białko-białko)

AGFK — mieszanina L-asparagianinu, D-glukozy, D-fruktozy i jonów potasów

MPa — mega Paskal

PFPP — motyw: prolina, fenyloalanina, prolina, prolina

PE — fosfotydyloetanonloamina

PG — fosfatydyloglicerol

CL — kardiolipina

DAG — diacyglicerol

PDBQT — Protein Data Bank, Partial Charge (Q), & Atom Type (T)

FFT — ang. *Fast Fourier Transform* (Szybka Transformata Fouriera)

TLC — ang. *Thin Layer Chromatography*

TM — ang. *Transmembrane*

APC — ang. amino acid/polyamine/organocation

In silico — z łac. w krzemie

Ab initio — z łac. od początku

In vivo — z łac. na żywym

In vitro — z łac. w szkle

3. WSTĘP TEORETYCZNY

In silico (łac. w krzemie) w odniesieniu do badań biologicznych oznacza badania, które zostały przeprowadzone z wykorzystaniem komputera i technik informatycznych.

Modelowanie molekularne jest dziedziną nauki, która poprzez budowanie modeli i wykorzystywanie obliczeń komputerowych zajmuje się przewidywaniem struktur molekularnych i ich funkcji. Kluczowym w modelowaniu molekularnym jest zatem opracowywanie oraz stosowanie modeli odpowiednio dostosowanych do zagadnień, które są za ich pomocą badane i rozwijane.

Dziedzina nauki zajmująca się modelowaniem molekularnym ma swoje początki w latach 60-tych XX wieku, lecz jej rozwój nabrał rozpędu dopiero w połowie lat 80-tych XX wieku wraz z pojawieniem się pierwszych superkomputerów. Rozwój, napędzany postęпами w takich obszarach jak m.in. genomika, grafika komputerowa, algorytmy i bazy struktur białek, w połączeniu z rosnącą mocą obliczeniową oraz pamięcią komputerów, sprawiły, że modelowanie teoretyczne jest aktualnie metodą szeroko wykorzystywaną w badaniach eksperymentalnych (Schlick, 2010).

3.1. Struktura białek

W architekturze białek wyróżnia się cztery poziomy organizacji ich struktury.

Pierwszy poziom, struktura pierwszorzędowa białek, to sekwencja (ciąg) reszt aminokwasowych budujących konkretne białko. Do struktury pierwszorzędowej czasami zaliczane jest też umiejscowienie mostków disulfidowych. Struktura drugorzędowa opisuje wzajemne (wysoce uporządkowane i lokalnie powtarzalne) przestrzenne ułożenie sąsiadujących ze sobą (w sekwencji liniowej) reszt aminokwasowych. Przykładami struktury drugorzędowej są α -helisy i β -kartki.

Struktura trzeciorzędowa opisuje relacje przestrzenne i wzajemne ułożenie reszt aminokwasowych, które są od siebie oddalone w sekwencji liniowej. Ponadto, struktura trzeciorzędowa opisuje również lokalizację mostków disulfidowych i innych oddziaływań stabilizujących strukturę przestrzenną białka. Struktura trzeciorzędowa jest najbardziej zachowywanym w ewolucji poziomem organizacji struktury białek.

Czwartym poziomem organizacji, strukturą czwartorzędową, nazywamy kompleks zbudowany z wielu łańcuchów białka (podjednostek). Tak więc, struktura czwartorzędowa odnosi się do wzajemnego ułożenia podjednostek białka w przestrzeni oraz natury ich oddziaływań. Najprostszym przykładem struktury czwartorzędowej mogą być dimery (Berg, Tymoczko i Stryer, 2005). Warto zaznaczyć, że nawet wśród białek homologicznych, struktura czwartorzędowa nie jest tak konserwowana w ewolucji jak struktura trzeciorzędowa (Nakamura, Oda, Fukasawa i Tomii, 2017).

3.1.1. Rola struktur i ich znaczenie

Funkcja białka jest bezpośrednio zależna od jego struktury trzeciorzędowej (Voet, Voet i Pratt, 2016). Tak więc, zdolność do pełnienia swoich funkcji biologicznych wymaga od białka przyjęcia specyficznej konformacji przestrzennej w spontanicznym lub wspomaganym procesie nazywanym zwijaniem (ang. *foldng*). Proces ten napędzany jest poprzez szereg oddziaływań niekowalencyjnych — wiązania wodorowe, oddziaływania hydrofobowe oraz siły Van der Waalsa. Oddziaływania hydrofobowe wraz z mostkami disulfidowymi (powstającymi pomiędzy dwiema resztami cysteiny) stabilizują zwiniętą strukturę. Zatem, zrozumienie funkcji białek na poziomie molekularnym często wymaga wcześniejszego określenia ich trójwymiarowej struktury.

Mutacje, które skutkują zaburzeniem struktury białka mogą prowadzić do zmian w jego funkcji. Zmiany te mogą być neutralne, szkodliwe lub korzystne dla funkcji białka, dlatego zrozumienie struktury białek jest niezbędne do pełnego zrozumienia ich biologii. Co więcej, struktury trzeciorzędowe dają wgląd w istotne cechy ewolucyjne całych rodzin białek (Jabeen, Mohamedali i Ranganathan, 2018).

Istnieje szereg metod eksperymentalnych (np. krystalografia rentgenowska (Kendrew i inni, 1958), NMR — jądrowy rezonans magnetyczny (Cavanagh, Fairbrother i Palmer, 2007)) pozwalających na wyznaczenie struktury białek oraz poznanie ich właściwości. Metody te są jednak kosztowne i czasochłonne. Struktury trzeciorzędowe, obok metod eksperymentalnych, mogą być przewidywane przy użyciu mniej kosztownego podejścia obliczeniowego (Khor, Tye, Lim i Choong, 2015). Różnica pomiędzy ilością sekwencji białek dostępnych w bazach danych a odpowiadającymi im strukturami białek drastycznie wzrasta, prowadząc do tego, że liczba dostępnych sekwencji białek wielokrotnie przewyższa ilość znanych nam struktur. Zjawisko to działa na niekorzyść badaczy, gdyż

prawdopodobnie dla znacznej większości białek nie będziemy w stanie określić struktury za pomocą metod eksperymentalnych. Dlatego też jednym z najważniejszych narzędzi, które pozwoli zmniejszyć tę różnicę w ilości sekwencji i struktur jest teoretyczne modelowanie struktur białek (Larsson, Wallner, Lindahl i Elofsson, 2008).

3.1.2. Teoretyczne metody przewidywania struktur białek

Modelowanie struktury trzeciorzędowej białka skupia się na przewidywaniu struktury na podstawie jego sekwencji aminokwasowej. Przewidywanie to odbywa się za pomocą algorytmów obliczeniowych. Podejścia obliczeniowe do przewidywania struktur białkowych można zasadniczo podzielić na trzy kategorie (modelowanie porównawcze, przewlekanie oraz podejście *ab initio* – (łac. od początku), każda z tych metod oparta jest na prawach fizyki oraz ewolucji (Khor, Tye, Lim i Choong, 2015; Illergård, Ardell i Elofs, 2009).

Modelowanie porównawcze

Modelowanie homologiczne jest najbardziej skutecznym podejściem w przewidywaniu struktur trzeciorzędowych. Opiera się ono na obserwacji, że struktura jest bardziej zachowywana w toku ewolucji niż sekwencja (Bajaj i Blundell, 1984). W związku z tym, podejście to wymaga znalezienia istniejącej struktury, która służyć będzie jako *szablon* (białko o znanej strukturze) w modelowaniu białka docelowego (białko, którego struktury poszukujemy). Najważniejsza zasada modelowania porównawczego mówi o tym, że im lepiej dobrany jest szablon, tym dokładniejsze są przewidywania (Fiser, 2010). To podejście zostało szerzej opisane w punkcie 3.2.

Przewlekanie

Przewlekanie (ang. *Threading*) jest metodą rozpoznawania pofałdowania (Jones, Taylor i Thornton, 1992) i opiera się na obserwacji, że w naturze istnieje określona liczba pofałdowań, które może przyjąć białko, dzięki temu nawet odległe ewolucyjnie białka homologiczne mogą mieć podobną strukturę trzeciorzędową (Kelley i Sternberg, 2009).

Modelowanie *ab initio*

Jest to modelowanie struktury białka na podstawie oddziaływań międzyatomowych. Opiera się na hipotezie termodynamicznego zwijania białek, według której natywna struktura białka odpowiada globalnemu minimum energetycznemu (Anfinsen, 1973). Modelowanie *ab initio* jest teoretycznie możliwe, aczkolwiek wymaga intensywnych i kosztownych obliczeń, ponieważ nawet niewielkie zmiany konformacyjne występujące w naturze wymagają długotrwałych obliczeń i symulacji. Jest najmniej dokładnym z trzech prezentowanych podejść oraz ograniczone jest rozmiarem białka, którego strukturę przewiduje. Dlatego też metody *ab initio* używa się do modelowania małych cząsteczek oraz niewielkich białek/peptydów (Lee, Freddolino i Zhang, 2017).

3.2. Modelowanie homologiczne

Głównym celem modelowania homologicznego, nazywanego też modelowaniem porównawczym, jest utworzenie modelu struktury białka na podstawie jego sekwencji. W idealnym przypadku dokładność takiego modelu powinna być zbliżona do tej uzyskiwanej za pomocą eksperymentalnych metod ustalania struktur przestrzennych białek.

Modelowanie porównawcze oparte jest na podobieństwie sekwencji białka docelowego, czyli białka, którego strukturę pragniemy poznać, do białka homologicznego o znanej strukturze dostępnej w bazie PDB (Greer, 1981; Fiser, 2010) i można je opisać w czterech głównych krokach:

- i. poszukiwanie szablonu, czyli białka homologicznego na podstawie którego zbudowany zostanie model białka docelowego. Poszukiwanie homologa wymaga identyfikowania podobieństw w sekwencji białka docelowego i sekwencji białek o znanych strukturach
- ii. zestawienie (ang. *alignment*) sekwencji białka docelowego z sekwencją/sekwencjami szablonu/szablonów oraz ewentualne korekcje zestawienia
- iii. budowanie modelu opartego na zestawieniu sekwencji białka docelowego z wybranymi szablonami oraz na strukturze szablonów
- iv. ocena jakości modelu oraz poprawa błędów modelowania.

Budowanie modelu (krok iii) jest procesem skomplikowanym, składającym się z kilku etapów. W pierwszym etapie budowany jest łańcuch główny modelu białka. Następnie modelowane są łańcuchy boczne oraz pętle (dodatkową optymalizację pętli można wykonać po uzyskaniu gotowego modelu). Proces kończy się optymalizacją modelu. Oprogramowania służące do modelowania homologicznego są dostępne w formie aplikacji desktopowych jak również w formie serwerów on-line i mają wbudowane algorytmy optymalizacji, np. MODELLER (Sali i Blundell, 1993).

Tak przygotowany model podlega następnie ocenie jakości i weryfikacji poprawności. Narzędzia służące do modelowania zawierają szereg funkcji oceniających, na podstawie których jesteśmy w stanie wybrać najlepszy z uzyskanych modeli. Wybrane w ten sposób modele należy poddać pogłębionej ocenie jakości (np. obliczając wartość RMSD; rozdział 3.2.2).

3.2.1. Modelowanie na wielu szablonach

Tradycyjnie w metodach modelowania porównawczego używa się jednego szablonu, który wykorzystuje się do modelowania struktury białka docelowego. Taki szablon powinien być możliwie jak najlepiej oceniany w trakcie poszukiwania homologów (oceniane jest m.in. podobieństwo sekwencyjne obu białek). Niestety nie zawsze możliwe jest wybranie jednoznacznie najlepszego szablonu (Venclovas i Margelevicius, 2005). Wynika to z różnych przyczyn, m.in. białko docelowe może nie posiadać jednego bliskiego homologa, na którego podstawie można byłoby przewidzieć jego strukturę.

W takim przypadku można skorzystać z metody modelowania homologicznego opartej na wielu szablonach (ang. *multitemplate homology modeling*). Metodę modelowania wieloszablonowego po raz pierwszy zaprezentowano w eksperymencie CASP7 w 2006 roku. Podczas eksperymentu użyto kilku serwerów stosowanych do automatycznego modelowania homologicznego i przeprowadzono modelowania oparte na wielu szablonach w celu poprawienia jakości przewidywanych modeli. Część z wykorzystanych w eksperymencie serwerów (np. Zhang-server) zbudowało w ten sposób modele o zadowalających rezultatach i jakości (Cheng, 2008).

Modelowanie homologiczne z wykorzystaniem wielu szablonów podlega takim samym zasadom jak modelowanie homologiczne przeprowadzone na jednym szablonie.

3.2.2. RMSD jako wyznacznik podobieństwa strukturalnego

Jedną z najważniejszych i najczęściej stosowanych miar oceny jakości modelu jest wartość odchylenia średniej kwadratowej pozycji atomów — RMSD (ang. *root-mean-square deviation*). Obliczenie wartości RMSD opiera się na porównaniu dwóch struktur, wymaga więc nałożenia na siebie struktury modelu oraz struktury szablonu, na podstawie którego przeprowadzono przewidywanie struktury modelu. Nałożenie na siebie struktur nazywane jest superpozycją (ang. *superpositioning*) i polega na zastosowaniu przesunięć i rotacji jednej ze struktur tak, aby obie struktury posiadały optymalne dla porównania współrzędne. Następnie oblicza się pierwiastek z sumy kwadratów różnic między dwoma zbiorami atomów nałożonych na siebie struktur. Tak obliczone RMSD pokazuje, jak bardzo dwie struktury różnią się od siebie (Marti-Renom, Capriotti, Shindyalov i Bourne, 2009)

RMSD w dużym stopniu zależy od dokładności nałożenia na siebie dwóch struktur i jest silnie uwarunkowane tymi fragmentami struktur, które najbardziej się od siebie różnią (Kufareva i Abagyan, 2012). Im wyższa jest wartość RMSD, tym mniejsze jest podobieństwo porównywanych struktur. Dla modelowania homologicznego przyjmuje się, że wartości średniej odległości pomiędzy atomami w modelu i szablonie poniżej 2Å oznaczają model podobny do szablonu, co implikuje, iż model ma dobrą jakość i jest wiarygodny.

RMSD jest najprostszą metodą oceny jakości modelu i pozwala szybko ocenić jakość przewidzianego modelu. Dobrą alternatywą dla RMSD jest obliczenie wartości globalnych odległości GDT (ang. *global distance test* (Read i Chavali, 2007)). Ta miara jakości modelu sprawdzi się, gdy przewidziana struktura zawiera wiele nieuporządkowanych pętli, które mogą odstawać podczas nakładania struktur i mieć znaczący wpływ na wartości RMSD. Poza obliczeniem wartości RMSD i/lub GDT, należy również poddać model dodatkowym ocenom, stosując inne metody i wykorzystując dostępne serwery lub oprogramowanie. Przykładowe strategie oceny jakości modelu oraz dedykowane im oprogramowania opisałam w podrozdziale 3.2.3.

3.2.3. Strategie oceny jakości modelu

Niezależnie od tego, czy model struktury białka został opracowany za pomocą metod modelowania homologicznego, rozpoznawania pofałdowania lub metod *ab initio*, modele takie powinny zostać poddane ocenie jakości z wykorzystaniem dedykowanych programów (MQAP — ang. *Model Quality Assessment Program*). Programy służące do tzw. walidacji jakości modelu dokonują oceny modelu odzwierciedlającej jego wiarygodność (Kryshtafovych i inni, 2016). Zazwyczaj taka ocena uwzględnia to, w jakim stopniu oceniany model teoretyczny struktury białka przypomina strukturę białka wyznaczoną metodami eksperymentalnymi (np. za pomocą krystalografii rentgenowskiej).

Ocena taka może opierać się na podobnych założeniach jak metody walidacji struktur eksperymentalnych lub może być specyficzna dla metod opartych na teoretycznym przewidywaniu struktur. Przykładowo może sprawdzać, czy liczba poszczególnych struktur drugorzędowych w modelu zachowuje relacje ilościowe obecne w znanych strukturach białek. Inne sposoby mogą opierać się np. na sprawdzaniu liczby grup hydrofobowych (Feenstra, Abeln i Heringa, 2018).

Przykładami programów służących do oceny struktur białek, dedykowanych dla oceny modeli teoretycznych są QMBRANE (Studer, Biasini i Schwede, 2014) oraz QMEANDisCo (Studer G. i inni, 2020) dostępne na serwerach SWISS-MODEL (QMEAN) lub serwer PSVS zawierający w sobie program PROCHECK (PSVS; Yuanpeng, Powers i Montelione, 2005).

3.3. Oddziaływania białek i dokowanie

Białka posiadają zdolność do oddziaływania z innymi makromolekułami (w tym również z innymi białkami) i szacuje się, że więcej niż 80% znanych białek wymaga interakcji z innymi cząsteczkami, by móc pełnić swoje funkcje (Berggård, Linse i James, 2007). Ponadto, wiele funkcji fizjologicznych wymaga formowania się kompleksów receptor-ligand (dokładniej mogą to być kompleksy: białko-ligand i/lub białko-białko). Badanie wzajemnych oddziaływań białek pokazuje nam także, jak białka oddziałują ze sobą w szlakach metabolicznych i sygnalizacyjnych (Keskin, Tuncbag i Gursoy, 2016).

Dokowanie jest metodą obliczeniową pozwalającą opisać te oddziaływania i skupia się na odnalezieniu orientacji, w jakiej wiążą się do siebie dwie molekuly o znanych strukturach. Dokowanie zazwyczaj obejmuje dwa powiązane ze sobą aspekty: (1) próbkowanie stanów konformacyjnych układu receptor-ligand, gdzie jedną cząsteczkę traktuje się jako ligand, a drugą jako receptor; (2) obliczenie oceny punktowej jako energii wiązania kompleksu lub innej oceny, która opisuje biologiczną aktywność uzyskanego kompleksu. Próbkowanie jako wynik zwraca konformacje kompleksu receptor-ligand. Oceny punktowe natomiast określają powinowactwo receptora do wiązania liganda (Novikov i Chilov, 2009).

Dokowanie, w przeciwieństwie do czasochłonnych i drogich metod eksperymentalnych, pozwala na sprawdzenie i ocenę tysięcy potencjalnych konformacji i tysięcy różnych ligandów stosunkowo niskim kosztem. Podobnie, w przypadku badania oddziaływań białko-białko, wykorzystanie dokowania molekularnego może dostarczyć wielu cennych informacji, np. o resztach aminokwasowych biorących udział we wiązaniu lub pozwala wyznaczyć potencjalne miejsca interakcji, co na ogół ułatwia przeprowadzenie dalszych eksperymentów. Od czasu pionierskiej pracy Kuntza i jego współpracowników (Kuntz, Blaney, Oatley, Langridge i Ferrin, 1982), w której po raz pierwszy opisano metodę badania geometrycznie możliwych dopasowań ligandów i receptorów, nastąpił znaczny postęp technologiczny, który pozwolił na poprawę szybkości i precyzji obliczeń (Huang i Zou, 2010).

3.3.1. Dokowanie białko-ligand

Dokowanie w układzie białko-ligand jest narzędziem pomagającym w prawidłowym zrozumieniu interakcji zachodzących między białkami i ich ligandami. Ten typ dokowania jest między innymi często używany na różnych etapach badań nad projektowaniem leków. W dokowaniu tego typu wymagane są dwie struktury: cząsteczka liganda oraz receptor białkowy, będące przedmiotem zainteresowania badania. Warto zaznaczyć, że ligandy są na ogół cząsteczkami o małej masie molekularnej (Elokely i Doerksen, 2013).

Dokowanie liganda do białkowego receptora pozwala przewidywać sposoby wiązania liganda do receptora w warunkach równowagi w roztworze wodnym. Dodatkowo, dzięki dalszym badaniom nad kompleksem białko-ligand, można wytypować fragmenty receptora odpowiedzialne za oddziaływanie z ligandem, niezależnie czy są to całe regiony

białka, czy pojedyncze reszty aminokwasowe, co często może służyć zidentyfikowaniu reszt potencjalnie istotnych dla funkcjonowania badanego białka. Trzema głównymi aspektami algorytmów opisujących dokowanie liganda do receptora białkowego są: próbkowanie konformacji liganda, elastyczność białka oraz odpowiednio dobrana funkcja oceniająca.

Wiązanie liganda do białka zazwyczaj skutkuje zmianami konformacyjnymi białka (tzw. dopasowanie indukowane, z ang. *induced fit*). Zmiany te mogą być niewielkie i dotyczyć jedynie rearanżacji przestrzennej łańcuchów bocznych białka w jego kieszeni wiążącej. Związanie liganda może wiązać się jednak także z ruchami całych domen receptora. Ze względu na rozmiary i skomplikowanie struktur białek (wynikające z wielu stopni swobody), elastyczność receptora jest najtrudniejszym z wyzwań, z jakimi mierzą się programy dokujące (Huang i Zou, 2010). W związku z tym dostępne metody dokowania różnią się od siebie w zakresie tego, w jakim stopniu zezwalają na elastyczność białka.

Kolejnym istotnym problemem związanym z dokowaniem liganda do receptora jest znalezienie miejsca wiążącego receptora oraz określenie konkretnej orientacji liganda w tym miejscu. Zdarza się, że lokalizacja kieszeni wiążącej białka jest nieznana. W takiej sytuacji pierwszym krokiem do przeprowadzenia poprawnych obliczeń jest wytypowanie potencjalnych miejsc wiążących receptora. Dokowanie globalne (ang. *global docking*) pozwala na przeprowadzenie próbkowania liganda na całej powierzchni białka, dzięki czemu umożliwia identyfikację miejsca wiązania dla wybranej cząsteczki wraz ze wskazaniem korzystnych konformacji liganda. Innym sposobem identyfikacji miejsc wiążących może być skorzystanie z metod takich jak np. wykrywanie zagłębień w powierzchni białka. W obu przypadkach, po zidentyfikowaniu kieszeni wiążących warto dodatkowo lokalnie zadokować do niej ligand.

Dokowanie lokalne (ang. *local docking*) polega na próbkowaniu konformacji liganda i wiązaniu go do receptora w określonym obszarze, który zidentyfikowany jest jako powierzchnia wiążąca białka. Jest to podejście wymagające wiedzy na temat miejsc wiążących receptora. Informacje tego typu mogą wynikać z przeprowadzonych eksperymentów biochemicznych lub z informacji o kieszeniach wiążących struktur białek homologicznych do białka badanego. Dokowanie lokalne jest mniej kosztowne obliczeniowo niż dokowanie globalne, a jego wyniki są dokładniejsze.

3.3.2. Dokowanie białko-białko

Interakcje białko-białko (PPI ang. *protein-protein interactions*) mogą powodować utworzenie stabilnych lub przejściowych kompleksów. Białka oddziałują ze sobą w sposób przejściowy, gdy celem określonej interakcji jest uzyskanie efektu w krótkim czasie — ten typ interakcji dominuje w przekazywaniu sygnałów. W interakcjach przejściowych białka oddziałują tymczasowo i dysocjują *in vivo* (łac. na żywym). Stałe interakcje natomiast są bardzo stabilne, tworząc często permanentne kompleksy. Białka oddziałujące ze sobą w „nie przejściowy” sposób prawdopodobnie będą wspólnie ulegać ekspresji i/lub będą zlokalizowane w tym samym miejscu komórki (Nooren i Thornton, 2003).

Oddziaływania białko-białko można podzielić ze względu na efekt tych oddziaływań. Białka mogą budować homokompleksy lub heterokompleksy. Homokompleksy zbudowane są z takich samych podjednostek i najprostszym przykładem jest homodimer, czyli dimer złożony z dwóch identycznych podjednostek. Heterokompleksy zbudowane są przez różne podjednostki białkowe (Jones i Thornton, 1996). Homooligomery najczęściej są symetryczne i bardziej stabilne niż heterokompleksy (Keskin, Tuncbag i Gursoy, 2016).

Metody dokowania białko-białko mają wiele wspólnego z dokowaniem małych ligandów — wykorzystują tę samą kombinację algorytmów próbkujących i funkcji oceniających. Główna różnica (a zarazem największe wyzwanie) polega na tym, że przeszukiwana przestrzeń konformacyjna w przypadku wiązania do siebie dwóch (lub więcej) białek jest znacznie większa w porównaniu z tą, którą przeszukujemy w dokowaniu ligand-białko.

W przypadku dokowania białko-białko, podobnie jak dla dokowania ligand-białko, możemy wyróżnić metody dokowania globalne oraz lokalne.

W dokowaniu globalnym zazwyczaj jedna cząsteczka białka jest nieruchoma i nazywana jest ona receptorem, podczas gdy drugie białko, nazywane ligandem, jest obracane wokół receptora w celu znalezienia odpowiedniej wzajemnej orientacji. Dokowanie globalne kończy się w momencie, w którym uwzględnione zostały wszystkie możliwe orientacje białek w przestrzeni próbkowania. Takie podejście jest bardzo kosztowne pod względem obliczeniowym. Przykładem sposobu na redukcję kosztowności dokowania globalnego jest wykorzystanie dyskretnej transformacji Fouriera, powszechnie używanej do kompresji danych.

Dokowanie lokalne wykorzystuje informacje o strukturze powierzchni białek takie jak np. płaskość powierzchni czy regiony hydrofobowe. Takie cechy obu białek są następnie do siebie dopasowywane w trakcie dokowania w celu uzyskania odpowiedniej komplementarności pomiędzy nimi. W tym podejściu wykorzystuje się jedynie informacje o zasięgu lokalnym względem całej powierzchni białek (Keskin, Tuncbag i Gursoy, 2016).

W obu podejściach można wykorzystać informacje biochemiczne, biofizyczne lub chemiczne na temat miejsc i rodzaju interakcji zachodzących pomiędzy badanymi białkami, jeżeli takowe są znane. Dostępne dane pozwalają na zawężenie obszaru poszukiwań i zwiększenie dokładności obliczeń. W przypadku braku dostępnych danych doświadczalnych można skorzystać z komputerowych metod przewidywania interakcji białko-białko. Niektóre z tych metod opisuję szerzej w punkcie 3.3.3.

3.3.3. Obliczeniowe metody przewidywania regionów wiązania

Obok dokowania istnieją też inne metody przewidywania miejsc wiążących ligandy czy miejsc wiązania w interakcjach białko-białko. Mogą one być stosowane jako dopełnienie lub wstępne założenia metod dokowania molekularnego.

Metody przewidywania regionów wiązania bazujących na strukturze trójwymiarowej białek opierają się na tym, że większość małych ligandów dokuje się do zagłębień i wgłębień występujących na powierzchni białek (Sotriffer i Klebe, 2002). Jedną z popularnych metod jest metoda skupiona na poszukiwaniu specyficznych cech geometrycznych na powierzchni białka. Podstawową ideą metod opartych na pomiarach geometrii przestrzennej jest zlokalizowanie dużych zagłębień w strukturze białka, które mogłyby być kieszeniami wiążącymi ligandy.

Inne metody przewidywania miejsc wiązania białek oparte są, podobnie jak modelowanie homologiczne, na szablonach strukturalnych. Metody te polegają na wyszukaniu w coraz bogatszych bazach danych białek homologicznych, które mają znane kieszenie wiążące. Na podstawie zestawień sekwencji identyfikuje się regiony odpowiedzialne za wiązanie liganda w białku docelowym.

Rozwój technologii komputerowych umożliwił powstanie metod przewidujących miejsca wiązania opartych na rozwiązaniach sztucznej inteligencji, takich jak algorytmy uczenia maszynowego (ang. *machine-learning*) czy głębokiego uczenia (ang. *deep-learning*) (Zhao, Cao i Zhang, 2020).

3.4. Symbioza metod obliczeniowych i eksperymentalnych

W przypadku, gdy nie znamy struktury badanego białka modelowanie molekularne, dokowanie ligandów oraz inne techniki bioinformatyczne, pozwalają nam nie tylko uzyskać kontekst do interpretacji danych biochemicznych i biologicznych, ale także ułatwiają projektowanie oraz kierowanie eksperymentami. Z kolei informacje znalezione lub potwierdzone podczas badań laboratoryjnych pozwalają na poprawienie jakości oraz dokładności modeli komputerowych oraz ich kontekstów. Ze względu na silny związek między strukturą a funkcją molekuł modele teoretyczne w połączeniu z danymi doświadczalnymi mogą zbudować spójny obraz procesów, zjawisk czy interakcji.

Przykładem takiego hybrydowego działania, które łączy w sobie zarówno metody obliczeniowe jak i laboratoryjne jest opublikowany w 2015 roku protokół opisujący doświadczalne modelowanie molekularne kompleksów białko-białko z udziałem białek GPCR (Kufareva, Handel i Abagyan, 2015).

Symulacje przeprowadzane za pomocą metod *in silico* wspomagają również eksperymenty o wysokim stopniu skomplikowania poprzez mniej kosztowne sprawdzanie wstępnych hipotez lub modelowanie zmian zachodzących w układach cząsteczek, które są trudne do zaobserwowania w trakcie przeprowadzania eksperymentu *in vitro* (łac. w szkle). W ten sposób w 2020 Andoh i inni sprawdzili różnice we właściwościach fizykochemicznych pomiędzy dwoma listkami dwuwarstwy lipidowej (Andoh, Hayakawa i Okazaki, 2020).

3.5. *Bacillus subtilis*

Laseczka sienna (*Bacillus subtilis*) jest bakterią Gram-dodatnią, odkrytą i opisaną w 1835 roku przez niemieckiego uczonego Christiana Gottfrieda Ehrenberga. *B. subtilis* powszechnie występuje w glebie, spotyka się ją również w przewodzie pokarmowym ludzi i zwierząt. Należy do rodziny *Bacillaceae* (Cohn, 1872) i jest jedną z najlepiej poznanych bakterii, dzięki czemu jest używana jako modelowy organizm w badaniach nad bakteriami Gram-dodatnimi, a także m.in. do badań nad sporulacją (Errington, 2003), morfogenezą (Errington, 1993) czy organizacją dwuwarstwy lipidowej (Nickles i inni, 2017). *B. subtilis* był pierwszą Gram-dodatnią bakterią, której genom zsekwencjonowano (Kunst i inni, 1997).

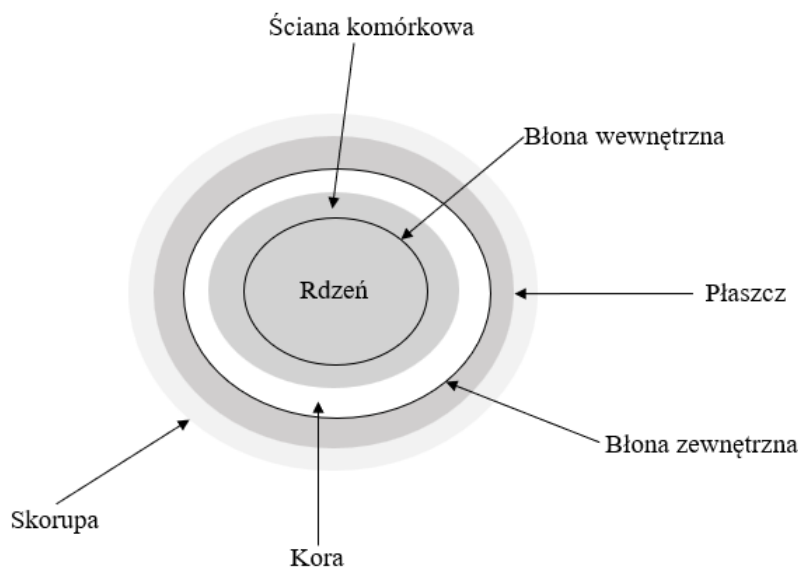
Bakterie z rodzaju *Bacillus* posiadają zdolność do tworzenia metabolicznie nieaktywnych form, zwanych przetrwalnikami lub sporami — tworzone są one, gdy środowisko zewnętrzne staje się niekorzystne dla wzrostu bakterii. Mechanizm ten pozwala bakterii przetrwać w niesprzyjających warunkach nawet przez wiele lat.

Przetrwalniki niektórych gatunków *Bacillus* są głównym czynnikiem odpowiedzialnym za psucie się żywności (*B. cereus*), stanowią również źródło chorób zwierząt i ludzi (*B. anthracis*), a także stanowią źródło zagrożenia bioterrorystycznego, przez co są przedmiotem niesłabnącego zainteresowania naukowców (Setlow i Johnson, 2012).

3.5.1. Budowa i właściwości przetrwalników

Przetrwalniki są nieaktywnymi metabolicznie formami komórki, które charakteryzują się wysoką odpornością na szkodliwe czynniki środowiskowe. Za takie czynniki uznaje się między innymi: brak wody, substancji odżywczych, wysoką i niską temperaturę, promieniowanie UV, a także niekorzystne pH. Przetrwalniki bakteryjne powstają w procesie zwanym sporulacją i charakteryzują się znacznym stopniem odwodnienia, co w połączeniu z ich specyficzną, wielowarstwową budową zapewniającą wysoką odporność na wyżej wymienione czynniki, pozwala bakterii przetrwać w takiej formie przez wiele lat bez utraty zdolności do kiełkowania, czyli powrotu komórki do formy wegetatywnej (Setlow P. , 2006; Setlow P. , 2014).

Rysunek 1 przedstawia uproszczony schemat budowy przetrwalnika *B. subtilis*. Najbardziej zewnętrzną warstwą spory jest skorupa (ang. *crust*). Pod skorupą znajduje się płaszcz, który otacza błonę zewnętrzną. Błona zewnętrzna umiejscowiona jest nad korą przetrwalnika zbudowaną ze zmodyfikowanego peptydoglikanu. Kolejną warstwą rozciągającą się pod ścianą komórkową jest wewnętrzna błona przetrwalnika. W tej błonie zlokalizowana jest większość białek i receptorów biorących udział w procesie kiełkowania przetrwalnika. Pod wewnętrzną błoną znajduje się rdzeń, który zawiera materiał genetyczny, rybosomy oraz większość enzymów. Rdzeń ma bardzo niską zawartość wody, zawiera za to duże ilości kwasu dipikolinowego (kwas pirydyno-2,6-dikarboksylowy, DPA, ang. *dipicolinic acid*), który uwalniany jest w procesie kiełkowania (Setlow P. , 2006).



Rysunek 1: Schematyczna budowa przetrwalnika *B. subtilis*. Zaadaptowano z (Setlow P. , 2014).

3.6. Kiełkowanie przetrwalników

Kiełkowanie przetrwalników (*germinacja*, ang. *germination*) jest nieodwracalnym procesem powrotu spor do formy wegetatywnej. Kiełkowanie można podzielić na dwie fazy, po których następuje etap zwany z ang. *outgrowth*. Rozpoczyna się ono niemal natychmiast po wykryciu w środowisku induktorów tego procesu (*germinantów*). Raz rozpoczęty, proces kiełkowania będzie postępować nawet po usunięciu *germinantów* ze środowiska (Paidhungat i Setlow, 2002).

Bakteryjne endospory muszą nieustannie monitorować środowisko zewnętrzne, by w odpowiednim momencie, gdy warunki środowiskowe staną się bardziej korzystne, zainicjować proces kiełkowania (*germinacji*). Sygnałem biochemicznym do rozpoczęcia *germinacji* może być wzrost stężenia specyficznych induktorów — *germinantów* — w środowisku spory oraz czynnik fizyczny jakim jest wysokie ciśnienie hydrostatyczne.

Inicjacja procesu kiełkowania indukowanego przez *germinanty* zachodzi poprzez receptory kiełkowania (RK) (Setlow P. , 2006; Setlow P. , 2014). Do tej pory scharakteryzowano trzy receptory kiełkowania bakterii *B. subtilis*: GerA, GerB oraz GerK, zlokalizowane w wewnętrznej błonie przetrwalnika. Wszystkie wymienione receptory związane są z odpowiedzią na różne induktory kiełkowania (Hudson i inni, 2001; Moir,

Kemp, Robinson i Corfe, 1994). Kiełkowanie Ger-A zależne indukowane jest przez interakcje z L-alaniną i/lub L-waliną, natomiast receptory GerB i GerK wspólnie inicjują kiełkowanie w obecności mieszaniny AGFK (L-asparaginan, D-glukoza, D-fruktoza oraz jony potasu) (Paidhungat i Setlow, 2000).

Jak wspomniano wyżej, inicjacja kiełkowania przetrwalników *Bacillus subtilis* może również zachodzić pod wpływem wysokiego ciśnienia (Gould i Sale, 1970). W zależności od wartości przyłożonego ciśnienia hydrostatycznego kiełkowanie rozpoczyna się poprzez aktywację receptorów kiełkowania (100-350 MPa) lub poprzez bezpośrednie uwalnianie jonów Ca^{2+} DPA (350-600 MPa) z pominięciem receptorów kiełkowania (Paidhungat M. i inni, 2002; Wuytack, Soons, Poschet i Michiels, 2000).

Proces kiełkowania przetrwalników *Bacillus subtilis* to proces, na który składa się kilka etapów. Pierwsze etapy skupione są wokół zmian zachodzących w samym przetrwalniku — pierwsze zmiany powodują uwodnienie rdzenia oraz utratę odporności przetrwalnika. Pojawienie się wody w strukturze przetrwalnika skutkuje degradacją kory oraz zmianami w obrębie dalszych warstw spory. Proces uwodnienia spory prowadzi ostatecznie do powrotu do formy wegetatywnej (etap *outgrowth*) oraz wznowienia metabolizmu (Powell i Strange, 1953; Moir, 2006; Setlow, Wahome i Setlow, 2008).

3.7. Receptory kiełkowania

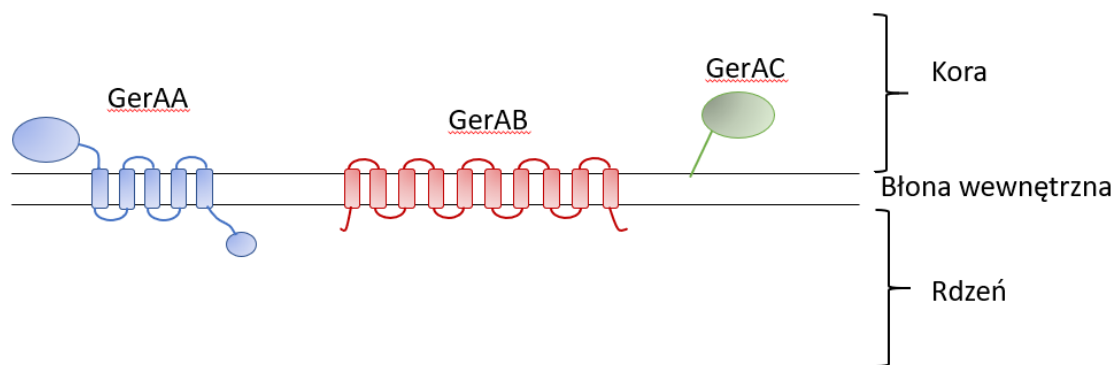
Receptory kiełkowania (RK) odpowiedzialne są za specyficzne inicjowanie procesu kiełkowania przetrwalników. RK zbudowane są z trzech podjednostek: A, B oraz C. Wszystkie podjednostki są niezbędne do ich prawidłowego funkcjonowania, a utrata którejkolwiek z nich skutkuje unieczynnieniem receptora.

Dodatkowo, poza receptorami kiełkowania, w procesie odpowiedzi na germinanty udział bierze białko GerD (Pelczar, Igarashi, Setlow i Setlow, 2007). GerD jest prawdopodobnie zakotwiczone w wewnętrznej błonie przetrwalnika i jest odpowiedzialne za poprawną klasteryzację receptorów. Receptory kiełkowania, wraz z lipoproteina GerD skupione są w jednym miejscu w wewnętrznej błonie przetrwalnika i miejsce to nazywane jest germinosomem. Germinosom z kolei swoją organizacją przypomina kompleks chemoreceptorów *E. coli* (Griffiths, Zhang, Cowan, Yu i Setlow, 2011; Korza i Setlow, 2013; Hudson i inni, 2001).

3.7.1. Receptor kiełkowania GerA

Receptor GerA cieszy się szerokim zainteresowaniem badaczy i pojawia się w wielu publikacjach związanych z procesem kiełkowania przetrwalników *B. subtilis*. Jednakże, pomimo obszernych badań, mechanizm działania, struktura oraz stechiometria podjednostek pozostają niewyjaśnione. Receptor GerA, podobnie jak inne receptory kiełkowania składa się z trzech podjednostek. Tworzą go białka GerAA, GerAB oraz GerAC. Sekwencja aminokwasowa białka GerAB wykazuje słabą homologię do podrodziny jednoskładnikowych błonowych transporterów bakterii, natomiast podjednostki A i C nie wykazują homologii sekwencyjnej z innymi opisanymi białkami (Moir, 2006; Jack, Paulsen i Saier, 2000). Wszystkie podjednostki są natomiast homologami odpowiednich podjednostek receptorów GerB i GerK, a także białek budujących receptory kiełkowania innych gatunków bakterii tworzących spory.

Jak już wspomniano w poprzednich akapitach, receptor GerA zlokalizowany jest w wewnętrznej błonie przetrwalnika, a dwie z jego podjednostek są integralnymi białkami błonowymi. Przewiduje się, że podjednostka GerAA, zbudowana z 482 reszt aminokwasowych, składa się z dwóch domen hydrofilowych: większej N-terminalnej i mniejszej C-terminalnej, oraz z domeny transmembranowej zbudowanej z 4-6 α -helis perforujących błonę (Wilson, Carlson, Janes i Hanna, 2012; Korza i Setlow, 2013). W ostatnim czasie określono strukturę rozpuszczalnej domeny białka homologicznego do GerAA (Li i inni, 2019). Podjednostka GerAB (365 aa) jest integralnym białkiem błonowym z przewidzianymi 10 helisami transmembranowymi oraz końcami (zarówno N-, jak i C-końcem) zlokalizowanymi w cytoplazmie (Cooper i Moir, 2011). Trzecia z podjednostek, lipoproteina GerAC (373 aa), nie jest integralnym białkiem błonowym, natomiast najprawdopodobniej jest przymocowana do błony za pomocą kotwicy diacyloglicerolowej połączonej z resztą cysteiny w pozycji 18. Badania pokazują, że taka kotwica jest często kluczowa dla poprawnego funkcjonowania receptorów kiełkowania (Vreeland, Rosenzweig i Powers, 2000).



Rysunek 2: Schemat budowy i rozmieszczenia podjednostek receptora GerA w wewnętrznej błonie przetrwalnika.

3.7.2. Interakcje receptora z germinantami

Receptor GerA indukuje kiełkowanie przetrwalników *B. subtilis* w obecności L-alaniny i/lub L-waliny. Według badań D-alanina jest silnym inhibitorem procesu kiełkowania rozpoczętego przez ten receptor. Wszystkie podjednostki receptora GerA są niezbędne do jego prawidłowego funkcjonowania, aczkolwiek funkcje poszczególnych podjednostek nie są poznane (Griffiths, Zhang, Cowan, Yu i Setlow, 2011).

Dowody na to, że białka receptora bezpośrednio oddziałują z germinantami pochodzą jedynie z badań dotyczących analizy wpływu różnych mutacji na funkcjonowanie tego receptora. Zbadany został wpływ substytucji aminokwasów w sekwencji białek GerAB (Cooper i Moir, 2011; Sammons, Moir i Smith, 1981) oraz GerAA (Mongkoltharuk, Cooper, Mawer, Allan i Moir, 2011) na zdolność przetrwalników do kiełkowania w obecności L-alaniny, a także na obecność podjednostek receptora GerA w germinosomie. Receptory, w których skład wchodziły mutanty białka GerAB z substytucjami L24F oraz T192I wymagały wyższego stężenia L-alaniny do rozpoczęcia procesu kiełkowania. Zmiany aminokwasów w bogatym w proliny motywie PFPP białka GerAA wykazały większą odpowiedź receptora na niskie stężenia L-alaniny i jej analogów, co wiązało się z szybszym inicjowaniem procesu kiełkowania w obecności niskiego stężenia germinantów, aczkolwiek w efekcie powrót do stanu wegetatywnego trwał dłużej niż w sporach typu dzikiego. Substytucje E330A oraz E335A podjednostki A skutkują brakiem kiełkowania w obecności L-alaniny, co związane jest ze zmniejszeniem ilości białka GerAC obecnego w receptorze o 95% (Moir, Lafferty i Smith, 1979).

Aktualny model procesu kiełkowania zakłada, że germinant przedostaje się do wewnętrznej błony spory, gdzie inicjuje nieznaną jeszcze mechanizm przesyłania sygnału, który rozpoczyna proces kiełkowania (Hudson i inni, 2001). Germinant wchodzi w interakcje z podjednostką A i B. Ewolucyjne pokrewieństwo białka GerAB z nadrodziną błonowych białek transportowych sugeruje, że może ono (GerAB) brać udział w wiązaniu liganda (Tharckay, Behravan, Southworth i Moir, 2001).

3.8. Germinosom

Receptory kiełkowania oraz białka GerD i SpoVA odgrywają istotną rolę we wczesnych etapach kiełkowania indukowanego germinantami i wszystkie te komponenty znajdują się w wewnętrznej błonie przetrwalnika. Badania pokazują, że poszczególne receptory kiełkowania współpracują ze sobą w trakcie indukowania procesu kiełkowania (Igarashi i Setlow, 2005). Biorąc pod uwagę ograniczoną mobilność podjednostek receptorów w wewnętrznej błonie przetrwalnika (Cowan, Koppel, Setlow i Setlow, 2003), uważa się, że białka receptorów mogą być zlokalizowane w jednym miejscu lub być zebrane w formie klastra jeszcze przed zainicjowaniem kiełkowania — przypominałoby to sposób organizacji receptorów chemotaksji wielu bakterii (Kentner i Sourjik, 2006). Białko GerD prawdopodobnie oddziałuje z RK, ułatwiając im skupienie się w jednym miejscu zwanym germinosomem.

3.8.1. Rola białka GerD w procesie kiełkowania

Mechanizm przekazywania sygnału inicjującego proces kiełkowania pomiędzy podjednostkami receptorów oraz pozostałymi składnikami procesu nie jest jeszcze poznany. Wyniki badań pokazują, że utrata białka GerD znacząco spowalnia proces kiełkowania przetrwalników bazujący na receptorach kiełkowania (Pelczar, Igarashi, Setlow i Setlow, 2007). Z badań nad strukturą i rolą białka GerD w procesie kiełkowania wyłania się model, według którego GerD jest białkowym rusztowaniem ułatwiającym skupianie się RK w germinosom. Ponadto GerD może również odgrywać rolę mediatora w procesie przekazywania sygnału inicjującego kiełkowanie.

Badania struktury białka GerD bakterii *Geobacillus stearothermophilus* pokazują, że formuje ono dimery, trimery i/lub kompleksy oligomeryczne wyższego rzędu. GerD jest prawdopodobnie białkiem peryferycznym, zakotwiczonym w wewnętrznej błonie przetrwalnika za pomocą kotwicy diacyloglicerolowej. Trimer GerD z kolei może samodzielnie tworzyć struktury wyższego rzędu i prawdopodobnie skupia się w jednym miejscu wewnętrznej błony (Li i inni, 2014; Pelczar i Setlow, 2008).

3.8.2. Organizacja receptorów w procesie chemotaksji

Chemotaksja jest systemem identyfikowanym u wielu gatunków bakterii m.in. u bakterii z rodzaju *Bacillus*. Chemotaksja określana jest jako ruch bakterii w kierunku zgodnym z rosnącym gradientem korzystnych czynników i/lub ucieczka od substancji szkodliwych. Tworzenie lub kiełkowanie spor są procesami indukowanymi poprzez wykrywanie braku lub obecności odpowiednich substancji w zewnętrznym środowisku bakterii. Oba systemy behawioralne opierają się zatem na monitorowaniu środowiska bakterii i reagowaniu na zachodzące w nim zmiany.

W procesie skupiania receptorów kiełkowania w geminosom znajdujący się w wewnętrznej błonie przetrwalnika dopatruje się szeregu podobieństw do przestrzennej organizacji chemoreceptorów, biorących udział w procesie chemotaksji (Griffiths, Zhang, Cowan, Yu i Setlow, 2011). Dokładne przyjrzenie się receptorom chemotaksji może dostarczyć odpowiedzi na pytania dotyczące organizacji receptora GerA w wewnętrznej błonie przetrwalnika oraz interakcji zachodzących między podjednostkami receptora, a także między receptorem a innymi białkami zaangażowanymi w inicjowanie procesu kiełkowania.

Badania pokazują, że receptory chemotaksji *E. coli* znajdują się w błonie komórkowej, a bliskie zależności pomiędzy receptorami odpowiadającymi na różne ligandy są istotne w szlaku sygnalizacyjnym (Shrout, Montefusco i Weis, 2003). Receptory chemotaksji bakteryjnej są niezwykle czułe na zmiany w chemicznym środowisku bakterii. Gromadzenie receptorów w jednym miejscu może być związane z ich wysoką wrażliwością na wspomniane wyżej zmiany w otoczeniu (Kim, Wang i Kim, 2002).

Wyniki badań sugerują, że wiązanie liganda do receptora chemotaksji nie wpływa na zmianę stanu oligomerycznego receptora w obrębie kompleksu sygnalizacyjnego. Tym samym podejrzewa się, że interakcje receptora z ligandem zachodzą w już stabilnym

kompleksie chemoreceptora. Taka obserwacja kontrastuje z mechanizmem działania innych systemów receptorowych, w których właśnie wiązanie liganda powoduje zmiany w strukturze czwartorzędowej oligomerów.

Minimalną jednostką strukturalną receptora chemotaksji jest symetryczny homodimer. Najczęściej taki dimer wiąże pojedynczą molekułę liganda. Wnioski płynące z kilku niezależnych badań pokazują, że tworzony przez dimery receptorów oligomer przyjmuje formę trimera, co stanowi rdzeń aktywnego kompleksu sygnalizacyjnego (Falke, 2002).

3.9. Błony lipidowe

Lipidy są zróżnicowaną grupą związków chemicznych spokrewnionych z kwasami tłuszczowymi. Lipidy pełnią trzy główne funkcje w organizmach żywych. Używane są jako magazyny energii — funkcjonują one jako magazyny rezerw kalorycznych, ale także dostarczają kwasów tłuszczowych oraz składników sterylowych do biogenezy błony. Z tego wynika druga funkcja lipidów — polarne lipidy formują macierz błony komórkowej, która składa się z części hydrofobowej i hydrofilowej. Dodatkowo lipidy mogą zachowywać się jak nośniki informacji w przekazywaniu sygnałów oraz w molekularnych procesach rozpoznawania innych cząsteczek (np. domeny błonowe wychytujące białka z cytozolu) (van Meer, Voelker i Feigenson, 2008).

Lipidy oraz białka są dwoma podstawowymi składnikami błon. Struktura błony jest wzmacniana przez szereg oddziaływań chemicznych. Obszary polarne błony zdominowane są przez oddziaływania elektrostatyczne, natomiast w obszarach niepolarnych dominują oddziaływania van der Waalsa oraz oddziaływania hydrofobowe.

W błonach lipidowych wyróżnić możemy trzy główne cechy charakterystyczne:

- i. płynność — składniki błony mogą znajdować się w ruchu,
- ii. asymetryczność — skład chemiczny błony może się różnić pomiędzy listkami,
- iii. heterogeniczność — różne rodzaje substancji w płaszczyźnie błony mają niejednorodny rozkład (Ślósarek, 2011).

3.9.1. Modelowanie błon fosfolipidowych

Techniki eksperymentalne mające na celu ustalenie organizacji błon stają się coraz bardziej zaawansowane. Jednakże, szczegółowa organizacja błon fosfolipidowych okazuje się trudna do zbadania na poziomie molekularnym przy użyciu, nawet zaawansowanych, technik eksperymentalnych ponieważ są to struktury bardzo dynamiczne, zbudowane z wielu zmiennych w czasie komponentów. Szczegółową informację o budowie błon mogą dostarczyć wyniki badań opartych na symulacjach komputerowych. Można tego dokonać przy użyciu technik takich jak dynamika molekularna (DM), która umożliwia opisanie interakcji pomiędzy wszystkimi komponentami symulowanego systemu na poziomie atomowym, jeżeli zapewni się im odpowiednią moc obliczeniową (Marrink i inni, 2019).

Pierwsze doniesienia o symulacji DM substancji powierzchniowo czynnych i lipidów ukazały się w latach 80-tych, krótko po pierwszej publikacji dotyczącej symulacji białek (Mccammon, Gelin i Karplus, 1977).

Zasadnicza trudność w zbudowaniu modelu kompleksu białko-błona wynika z problemu, jakim jest zanurzenie białka w modelu błony. Istnieją dwie metody powszechnie używane do budowania realistycznych kompleksów białko-błona.

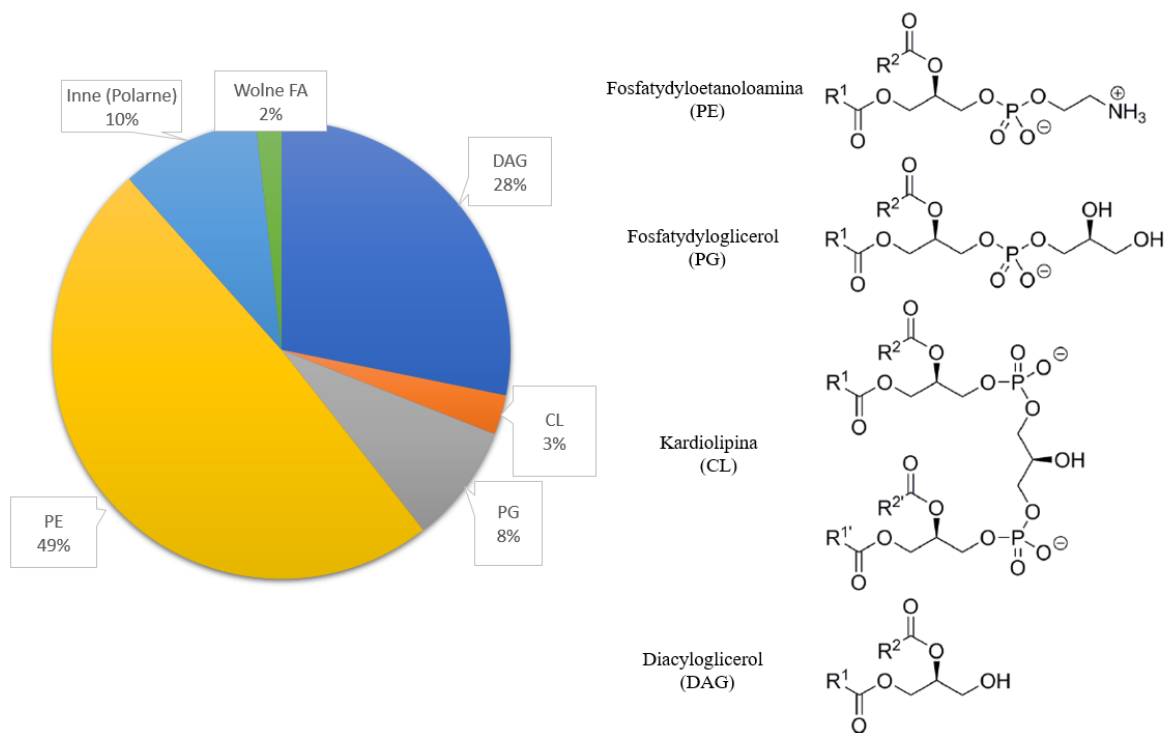
W pierwszej metodzie, podobne do lipidów pseudoatomy są rozprowadzone wokół białka, a następnie są pojedynczo zamieniane molekułami lipidów. Pojedyncze lipidy są wybierane z biblioteki lipidów dostępnej w wybranym oprogramowaniu, która zawiera różne konformacje cząsteczek lipidów. Ta metoda, nazywana *replacement method*, pozwala na łatwe kontrolowanie rozmiaru systemu i liczby lipidów.

Druga metoda, *insertion method*, polega na tym, że we wstępnie zrównoważonym modelu błony lipidowej tworzy się otwór, w którym następnie zanurza się białko błonowe. Ta metoda zapewnia dobrze zrównoważony model błony. Metoda ta wymaga mniej czasu oraz mocy obliczeniowej, jest jednak ograniczona rozmiarem budowanego systemu oraz kształtem białka — nadaje się dla białek o regularnych lub cylindrycznych kształtach (Jo, Kim i Im, 2007).

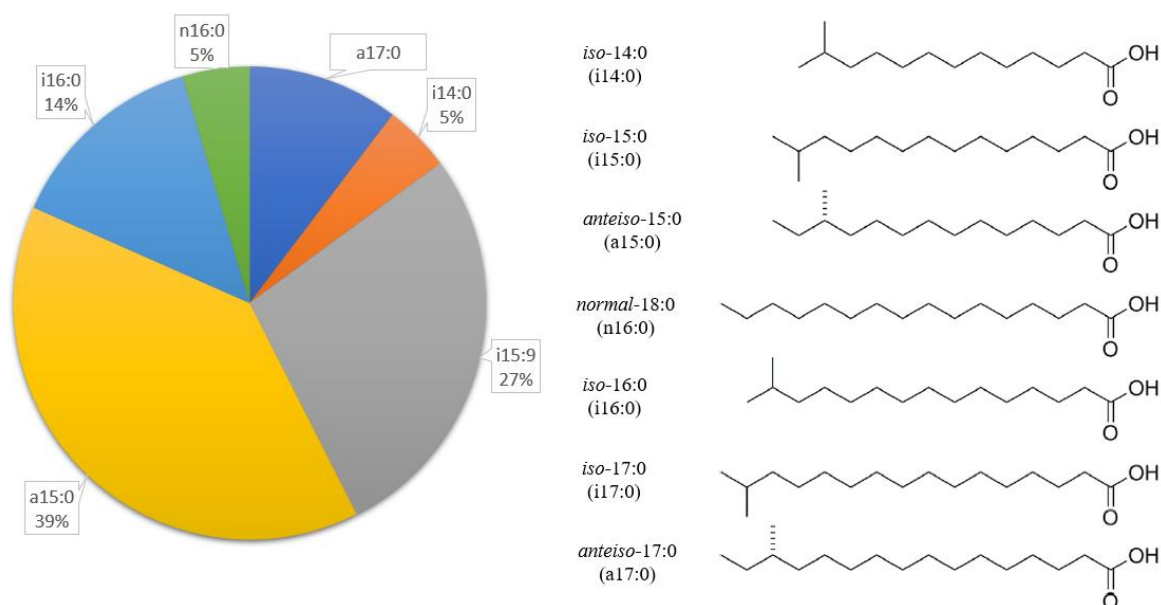
3.9.2. Model błon biologicznych *B. subtilis*

Ekstrakty lipidowe mogą dostarczyć informacji na temat kompozycji naturalnych dwuwarstw lipidowych, jednakże nie uwzględniają w pełni złożoności systemów, ponieważ nie zawierają informacji dotyczących istotnych kwestii jak organizacja czy asymetria warstw błony. Są one jednak dobrym punktem wyjściowym do badania składu błony i modeli. Modelowy ekstrakt lipidowy *B. subtilis* zawiera fosfatydyloetanolaminę (PE), fosfatydyloglicerol (PG), kardiolipiny (CL), lipidy lys-PG (lizylo-fosfatydyloglicerol) oraz około 30% neutralnych lipidów (głównie diacyloglicerol DAG) (Bishop, Rutberg i Samuelsson, 1967; Clejan, Krulwich, Mondrus i Seto-Toung, 1986; Den Kamp, Redai i Van Deenen, 1996; Kawai i inni, 2004).

Na podstawie tego ekstraktu zbudowany został model błony fosfolipidowej imitujący naturalne błony *B. subtilis*. Skonstruowano model zawierający 420 lipidów w każdej z warstw, o składzie bazującym na wspomnianym wyżej ekstrakcie lipidowym. Następnie, model poddano symulacjom dynamiki molekularnej przy użyciu oprogramowania GROMACS (Pronk i inni, 2013). Symulacje zostały przeprowadzone w czasie 1 μ s, w siedmiu stanach temperaturowych od 2 do 37°C w interwałach 5°C (Nickels i inni, 2017). Model ten jednak nie uwzględnia podziału na błony wewnętrzne i zewnętrzne. Niemniej, stanowi dobry punkt startowy w tworzeniu modelu błony wewnętrznej przetrwalnika *B. subtilis*. Dokładna kompozycja ekstraktu lipidowego opisana jest na wykresach zaprezentowanych na rysunku 3 oraz rysunku 4. Najczęściej w ekstrakcie występują fosfatydyloetanolaminy oraz diacyloglicerol. Obecne w ekstrakcie kwasy tłuszczowe zbudowane są z 14 do 17 atomów węgla, przy czym największy odsetek stanowią łańcuchy boczne o 15 atomach węgla zawierające rozgałęzienia typu *iso*- oraz *anteiso*-.



Rysunek 3. Po prawej stronie znajduje się wykres przedstawiający skład ekstraktu lipidowego *B. subtilis*: grupy główne. Po lewej — struktura grup głównych czterech lipidów najczęściej występujących w ekstrakcie. Na podstawie (Pronk i inni, 2013).



Rysunek 4. Po prawej stronie znajduje się wykres przedstawiający skład ekstraktu lipidowego *B. subtilis*: kwasy tłuszczowe. Wszystkie kwasy tłuszczowe są nasycone, głównie mają rozgałęzione łańcuchy i zawierają 14-17 atomów węgla. Po lewej — struktury siedmiu głównych kwasów tłuszczowych występujących w ekstrakcie. Na podstawie (Pronk i inni, 2013).

3.10. Znaczenie badań nad strukturą receptora GerA

Brak określonych struktur trzeciorzędowych białek GerAA, GerAB, GerAC budujących receptor kiełkowania GerA oraz białka GerD, które odpowiada za poprawne formowanie się germinosomu, prowadzi do istotnych braków w zrozumieniu mechanizmu działania oraz relacji funkcji i struktury tego receptora. Nieznane są tak istotne informacje jak przestrzenna budowa całego receptora oraz to, w jaki sposób podjednostki oddziałują ze sobą. Badania, w których podjęto próbę określenia liczby poszczególnych podjednostek białkowych receptora kiełkowania w przetrwalnikach pokazują, że stosunek stechiometryczny podjednostki A i C wynosi 1:1 (Steward i Setlow, 2013). Badania wskazują na zachodzenie interakcji pomiędzy podjednostkami A i B receptora (Moir, Kemp, Robinson i Corfe, 1994; Paidhungat i Setlow, 1999), inne badania sugerują również, że białko GerAC oddziałuje z białkami GerAA i GerAB (Igarashi i Setlow, 2005).

Zbudowanie modeli teoretycznych struktur białek receptora kiełkowania pozwoli na zrozumienie zależności między strukturą oraz funkcją receptora, a także na wyznaczenie potencjalnych miejsc wiązania ligandów. Modele struktury receptora i jego orientacji w modelu wewnętrznej błony przetrwalnika będą stanowiły kolejny element niezbędny w badaniach mających na celu wyjaśnienie mechanizmu działania receptora. Znajomość mechanizmu inicjacji procesu kiełkowania jest istotna w opracowaniu nowych metod umożliwiających eliminację przetrwalników z produktów spożywczych oraz minimalizujących zagrożenie związane z rozwojem chorób wywołanych przez przetrwalniki innych bakterii.

4. CEL PRACY

Celem niniejszej pracy badawczej było zbudowanie modeli teoretycznych białek wchodzących w skład receptora GerA oraz na ich podstawie, zbudowanie modelu receptora. Drugim celem było zbadanie potencjalnych miejsc wiążących germinanty w receptorze GerA.

5. MATERIAŁY I METODY

5.1. Przewidywanie struktur drugorzędowych

Przewidywanie struktur drugorzędowych każdego z badanych białek wykonałam z użyciem algorytmu JPred (Drozdetskiy, Cole, Procter i Barton, 2015). Wyniki przewidywania struktur drugorzędowych w dalszych częściach pracy szczegółowo porównałam z wynikami modelowania homologicznego struktur trzeciorzędowych.

5.2. Przewidywanie regionów transbłonowych

Przewidywanie regionów transbłonowych białek GerAA i GerAB przeprowadziłam z użyciem trzech niezależnych narzędzi: TMAP (Persson i Argos, 1994; Persson i Argos, 1996), TPred (Hofmann i Stoffel, 1993) oraz TMHMM (Sonnhammer, von Heijne i Krogh, 1998).

5.3. Wyszukiwanie homologów i zestawiania sekwencji

Szablony wykorzystane w modelowaniu homologicznym zostały przeze mnie znalezione przy pomocy serwera I-Tasser (Yang i inni, 2015; Roi, Kucukural i Zhang, 2010; Zhang, 2008). I-Tasser jest powszechnie wykorzystywanym narzędziem do przewidywania struktur oraz funkcji białek. W pierwszym etapie swoich przewidywań definiuje strukturalne szablony na podstawie zasobów bazy danych PDB oraz wielokrotnego nawlekania (ang. *multi threading*). Najlepiej oceniane struktury homologiczne, wykorzystane przez I-Tasser do przewidywania struktur podjednostek GerA, zostały wybrane jako szablony do modelowania molekularnego przeprowadzanego w ramach niniejszych badań.

Zestawienia sekwencji (PSA, ang. *pairwise sequence alignment*) wykonałam z użyciem algorytmu EMBOSS Matcher, który identyfikuje lokalne podobieństwa pomiędzy dwoma sekwencjami i jest oparty na aplikacji Laling Billa Pearsona wersji 2.0u4 z roku 1996. PSA dla globalnych podobieństw sekwencyjnych zostały wykonane z użyciem narzędzia EMBOSS Needle, które wykorzystuje algorytm Needleman-Wunsh (PSA). Wszystkie zestawienia zostały wykonane na podstawie macierzy BLOSUM62. Zestawienia

wielosekwencyjne (MSA, ang. *multiple sequence alignment*) dla modelowania na wielu szablonach dla podjednostki GerAB przeprowadziłam z użyciem algorytmu MUSCLE i/lub MAFFT (MSA).

5.4. Przewidywanie struktur trzeciorzędowych

Modelowanie homologiczne podjednostek receptora kiełkowania GerA oraz białka GerD wykonałam przy użyciu oprogramowania MODELLER w wersji 9.14. Modele teoretyczne obliczyłam z uwzględnieniem ograniczeń przestrzennych, na których opiera się modelowanie porównawcze wykonywane przez zaimplementowane w oprogramowaniu MODELLER algorytmy. Optymalizacja struktur oraz minimalizacja energii wykonałam przy użyciu domyślnych parametrów używanych w protokołach MODELLERA. Wygenerowane modele sklasyfikowałam na podstawie punktacji funkcji oceniających GA341 oraz DOPE, a następnie poddane dodatkowej ocenie jakości (Webb i Sali, 2016; Marti-Renom, Stuart, Fiser, Melo i Sali, 2000; Sali i Blundell, 1993).

Model białka GerAA budowałam dwukrotnie. Pierwszy model wykonałam przed publikacją struktury bliskiego homologa tego białka i użyłam modelowania wieloszablonowego. Szablony wybrałam przy pomocy serwera I-Tasser, do modelowania wykorzystano fragmenty szablonów, które wskazane zostały podczas globalnych i lokalnych zestawień sekwencji jako najbardziej podobne do sekwencji białka docelowego. Modelowanie przeprowadziłam na następujących szablonach: 3K07 (Long i inni, 2010) fragment 61-569), 4MT1 (Lei i inni, 2014) fragment 54-501) oraz 4DNT (Su i inni, 2012) fragment 59-555) z wymuszeniem utworzenia się helisy w regionie 272-302. Utworzenie helisy wymusiłam poprzez ustawienie więzów na ten fragment struktury podczas modelowania w programie MODELLER. Wszystkie białka użyte jako szablony są białkami błonowymi biorącymi udział w transporcie w poprzek błon biologicznych.

Drugi model białka GerAA, który wykorzystano do dalszych badań, zbudowałam po opublikowaniu struktury homologa tego białka. W tym przypadku do modelowania GerAA wykorzystałam dwa szablony: białko 6O59 (Li i inni, 2019) — podjednostka A receptora kiełkowania GerK *B. megaterium* — jako wzorzec dla domeny rozpuszczalnej oraz 3WDO (Jiang i inni, 2013) — *E. coli* YajR transporter — jako wzorzec integralnej domeny błonowej. Modelowanie przeprowadziłam w czterech etapach. W pierwszym kroku zbudowałam model części niebłonowej białka, następnie przewidziałam model domeny

blonowej na podstawie fragmentu białka 3WDO odpowiadającemu fragmentowi blonowemu GerAA. Oba modele następnie połączyłam za pomocą nałożenia na siebie wspólnego fragmentu struktury w programie PyMOL (Schrödinger, 2015). Tak „połączona” struktura białka posłużyła następnie jako szablon do przeprowadzenia modelowania dla całej sekwencji białka GerAA.

Model podjednostki B zbudowałam przy użyciu podejścia wieloszablonowego z czterema transporterami bakteryjnymi z rodziny APC jako szablonami: białka 3GIA, 3GI8 oraz 3GI9 (Shaffer, Goehring, Shankaranarayanan i Gouaux, 2009), które są niezależnymi od Na⁺ transporterami aminokwasów oraz białko 3L1L — antyporter argininy/agmatyny (Gao i inni, 2010).

Strukturę białka GerAC przewidziałam na podstawie bliskiego homologa strukturalnego — podjednostki B receptora kiełkowania GerB *B. subtilis* (Li, Setlow, Setlow i Hao, 2010).

Struktura białka GerD *Bacillus subtilis* przewidziałam na podstawie znanej struktury jego homologa — białka GerD organizmu *Geobacterium kaustophilus* (Li i inni, 2014). Homolog został skryształizowany w formie trimeru, dlatego modelowanie homologiczne przeprowadziłam zarówno dla trimeru GerD jak i monomeru. Struktura homologa została zdeponowana w bazie PDB pod kodem 4O8W.

5.5. Ocena jakości modeli teoretycznych

Wartość RMSD dla modeli w porównaniu do ich szablonów obliczałam na dwa sposoby: (1) przy wykorzystaniu kalkulatora wbudowanego w program Chimera (Pettersen i inni, 2004) oraz (2) przy pomocy serwera DALI (Holm i Sander, 1995; Holm, 2020). Dodatkowo, wszystkie modele oceniłam za pomocą wykresu Ramachandrana (Ramachandran, Ramakrishnan i Sasisekharan, 1963). Jakość modeli integralnych białek błonowych (GerAA i GerAB) ponadto potwierdziłam przy pomocy narzędzia QMBRANE, które jest oparte na wypadkowej potencjałów statystycznych ukierunkowanych na lokalne oszacowanie jakości modeli białek błonowych w ich naturalnym stanie (Studer, Biasini i Schwede, 2014). Białka GerAC oraz GerD dodatkowo ocenioniam przy użyciu funkcji oceniającej QMEANDisCo (Studer G. i inni, 2020).

5.6. Wizualizacja modeli

Wszystkie wizualizacje struktur białek przedstawiłam za pomocą oprogramowania Chimera oraz zaimplementowanych w nim funkcji pozwalających na zaznaczanie wybranych reszt aminokwasowych, fragmentów białka, obliczanie RMSD, ustawianie kolorów i oświetlenia modeli (Pettersen i inni, 2004). Modele kompleksów GerA w błonie przedstawiłam za pomocą oprogramowania PyMol (Schrödinger, 2015).

5.7. Dokowania molekularne

5.7.1. Wstępne analizy

Wstępna analiza potencjalnych miejsc wiążących ligandy dla białka GerAB oraz pierwszego modelu GerAA wykonałam za pomocą serwera MetaPocket. Serwer MetaPocket nie jest już wspieranym narzędziem. Twórcy tego serwera stworzyli nowe oprogramowanie dostępne online do przewidywań miejsc interakcji białko-ligand, jest to serwer *Protein-Ligand Interaction Profiler* (PLIP (Salentin, Schreiber, Haupt, Adasme i Schroeder, 2015)) i oba białka (podjednostka A oraz B) poddałam analizie w tym narzędziu. Analiza ta miała posłużyć weryfikacji oraz wsparciu wyników dokowania globalnego. Dodatkowo, w tym samym celu wykorzystałam serwer BiteNet powstały w 2020 roku (Kozlovskii i Popov, 2020).

Przewidywania oddziaływań białko-białko były pierwszym krokiem w kierunku utworzenia dimerów podjednostek. Do przewidywań wykorzystałam serwery meta-PPISP (Qin i Zhou, 2007) oraz cons-PPISP (Zhou i Shan, 2001). Serwer Cons PPISP bazuje na sieciach neuronowych, natomiast meta-PPISP jest serwerem zbudowanym na trzech różnych narzędziach (w tym cons-PPISP).

5.7.2. Dokowanie ligand-receptor

Dokowania globalne germinantów wykonałam przy użyciu oprogramowania AutoDock Vina. Jest to oprogramowanie dostępne na zasadach licencji *open source* i zostało, w swojej pierwotnej wersji, zaprojektowane i zaimplementowane przez doktora Olega Trott (Trott i Olson, 2010).

AutoDock Vina jest szeroko stosowanym oprogramowaniem do przeprowadzania eksperymentów dokowania molekularnego i jest znane ze swojej szybkości oraz dokładności. Oprogramowanie to wykorzystuje zaawansowane metody optymalizacji gradientowej w celu optymalizacji lokalnej kompleksu. Dodatkowo w oprogramowaniu zaimplementowano wielowątkowość, co pozwala na przyspieszenie jego działania i pracę na wielu rdzeniach procesora jednocześnie. W wyniku dokowania otrzymujemy nie tylko kompleksy, ale również wartość energii potrzebnej do wiązania i wartości RMSD ligandów (porównujące konformacje liganda przed i po zadokowaniu), które pokazują nam, jak bardzo zmieniła się konformacja liganda.

AutoDock Vina pracuje na plikach PDBQT, które wygenerowałam za pomocą narzędzia graficznego ADT MGLTools (Sanner, 1999) – rozszerzenie dodające interfejs GUI (ang. *graphical user interace*) dla programów AutoDock. To samo narzędzie posłużyło mi do ustalenia rozmiarów pudełka, czyli przestrzeni poszukiwań, w której ligand może się poruszać podczas dokowania. Do każdego białka przed rozpoczęciem dokowania dodałam polarne atomy wodoru, również za pomocą narzędzia ADT MGLTools.

Pierwszym krokiem w przeprowadzeniu eksperymentów dokowań dla białek podjednostek receptora GerA było dokowanie globalne. Ze względu na brak informacji o potencjalnych kieszeniach wiążących białek GerAA i GerAB (podjednostka C nie jest uznawana za oddziałującą z germinantami) dokowanie globalne miało za zadanie wskazać potencjalne miejsca, do których mogłyby się wiązać ligandy. Na podstawie wyników dokowania oraz liczebności klastrów wybrałam potencjalne kieszenie wiążące tego receptora (rozdział *Dokowanie germinantów*). AutoDock Vina pozwala na zadokowanie 20 kompleksów jednocześnie, w związku z tym dokowania zostały powtórzone wielokrotnie za pomocą krótkiego skryptu (załącznik 1), dzięki temu uzyskałam 20 · 2250 dla jednego powtórzenia, tych powtórzeń było od 1 do 5 w zależności od potrzeb. Uzyskane wyniki następnie przeanalizowałam w celu wybrania potencjalnych regionów wiążących germinanty.

Rozmiary przestrzeni poszukiwań (zwanej też pudełkiem) oraz jego koordynaty zostały spisałam w tabeli 1 dla każdego z dokowań globalnych. Do wszystkich białek przed dokowaniem dodane zostały atomy wodoru (również za pomocą narzędzia MGLTools).

W wyniku dokowań globalnych otrzymałam ogromne ilości danych i kompleksów, które należało uporządkować przed analizami. W celu uporządkowania wyników skorzystałam z zaimplementowanej w Gromacsie funkcji *gmx cluster* opartej na algorytmie opisanym przez Daura w 1999 roku (Daura i inni, 2004). *Gmx cluster* może grupować w

klastry struktury na podstawie odległości przy użyciu różnych metod. Odległości między strukturami mogą być obliczone na podstawie trajektorii, odczytane z pliku lub obliczone za pomocą RMSD. W swojej pracy korzystałam z metody opartej na obliczeniach RMSD z odcięciem dla dwóch sąsiadujących ze sobą struktur równym 0,1Å. Do dalszej analizy wybrałam klastry, w których liczebność kompleksów przekraczała 500.

Tabela 1: Rozmiary oraz koordynaty przestrzeni poszukiwań dla dokowań globalnych L-alaniny oraz L-waliny

Dokowanie globalne do monomeru GerAA		Dokowanie globalne do monomeru GerAB	
center_x	-19,164	center_x	35,631
center_y	82,007	center_y	11,959
center_z	71,95	center_z	4,145
size_x	66	size_x	70
size_y	108	size_y	52
size_z	48	size_z	60
Dokowanie globalne do dimeru GerAA		Dokowanie globalne do dimeru GerAB	
center_x	-6,083	center_x	0,0
center_y	79,472	center_y	0,0
center_z	53,556	center_z	3,39
size_x	112	size_x	126
size_y	110	size_y	126
size_z	86	size_z	100
Dokowanie globalne do kompleksu monomerów GerAA+GerAB		Dokowanie do kompleksu dimerów GerAA+GerAB	
center_x	-22,651	center_x	-2,711
center_y	47,062	center_y	70,191
center_z	62,287	center_z	56,266
size_x	70	size_x	108
size_y	126	size_y	126
size_z	78	size_z	86

5.7.3. Dokowania białko-białko

Dokowania białko-białko przeprowadziłam w celu utworzenia modeli dimerów, trimerów oraz modelu całego receptora. Dokowania przeprowadziłam przy użyciu następujących serwerów i programów dokujących: GRAMM-X, ClusPro i Sam.

Sam (ang. *Protein Structure Symmetry Assembler*) jest programem służącym do dokowania białek i bazuje na FFT. Sam został zaprojektowany do budowania symetrycznych kompleksów o dowolnej symetrii i dzięki wykorzystaniu FFT jest szybkim i skutecznym programem. Warty zaznaczenia jest to, że Sam działa bardzo dobrze dla monomerów o wielkości do 150 reszt aminokwasowych — przy większych monomerach dokowanie może być mniej dokładne (Ritchie i Grudinin, 2006). Sam został wykorzystany przeze mnie podczas budowania dimerów oraz trimerów wszystkich podjednostek. Najlepsze wyniki dawał dla podjednostki B. Wszystkie obliczenia wykonywane były na podstawie domyślnych ustawień programu z zaznaczeniem odpowiedniej symetrii — w przypadku dimerów była to symetria C2, a w przypadku trimerów symetria C3.

ClusPro jest serwerem, który od swojego powstania bardzo dobrze sobie radzi w eksperymentach CAPRI. Publikacje porównujące ten serwer z innymi znanymi metodami dokowań białko-białko uznają, że dysponuje on jednym z najlepszych algorytmów dokowania sztywnego oraz elastycznego (Desta, Porter, Xia, Kozakov i Vajda, 2020). Serwer ten wykonuje obliczenia w trzech etapach: 1) dokuje białka, używając metody korelacji FFT, 2) grupuje struktury w oparciu o RMSD w klastry oraz 3) udoskonala wybrane struktury (Vajda i inni, 2017; Kozakov i inni, 2017; Kozakov i inni, 2013). Serwer ten został wykorzystany do dokowań wszystkich podjednostek oraz, ze względu na to, że ClusPro korzysta z superkomputera do swoich obliczeń, do dokowania błonowej części kompleksu receptora GerA (zarówno zbudowanej na podstawie monomerów jak i dimerów). Podczas korzystania z serwera wykorzystywałam jego domyślnie ustawienia z dodatkowym oznaczeniem ilości monomerów biorących udział w dokowaniu, podczas budowania dimerów i trimerów.

GRAMM-X jest interfejsem sieciowym dostępnym dla oprogramowania GRAMM służącego do dokowania białek. Oba programy wykorzystują FFT do globalnego wyszukiwania najlepszych konformacji (Tovchigrechko i Vakser, 2006). Podobnie jak w przypadku ClusPro oraz Sam, serwer był wykorzystywany przeze mnie do budowania dimerów i trimerów. Na podstawie wyników z tego serwera wybrałam dimer GerAC.

GRAMM-X był przeze mnie wykorzystywany na domyślnych ustawieniach z wyjątkiem dimeru GerAC – na podstawie oddziaływań w dimerze homologa (GerBC) podałam potencjalnie reszty, które mogą brać udział w oddziaływaniach między monomerami GerAC.

Modele kompleksów poddałam takim samym ocenom jak modele monomerów podjednostek — skorzystałam z narzędzi QMbrane oraz QMEANDisCo, dodatkowo wybrane kompleksy oceniałam pod kątem energii dokowania oraz według funkcji oceniających wbudowanych w programy i algorytmy stosowane do zbudowania kompleksów.

5.7.4. Wizualizacja oddziaływań

Wizualizacja oddziaływań pomiędzy białkami w kompleksach oraz pomiędzy białkami i ligandami w niniejszej pracy przedstawiłam za pomocą narzędzia LigPlot+. Jest to program graficzny obsługujący programy LIGPLOT oraz DIMPLOT i generuje on interaktywne wykresy 2D pokazujące, jak ligand oddziałuje z receptorem — w bardzo przystępny sposób wyrysowane są wszystkie typy oddziaływań, długości wiązań i inne istotne informacje. LIGPLOT oblicza schematyczne wykresy oddziaływań pomiędzy białkiem i ligandem, a DIMPLOT z kolei, rozrysowuje oddziaływania białko-białko lub domena-domena (Laskowski i Swindells, 2011). LigPlot+ wykorzystałam w dalszych rozdziałach rozprawy zarówno do zobrazowania rozkładu oddziaływań między monomerami w kompleksach białek, jak i opisu oddziaływań podjednostek z germinantami.

5.8. Budowanie modelu błony fosfolipidowej

5.8.1. Badania eksperymentalne

Prof. dr hab. Antoni Banaś z Zakładu Biochemii Roślin Międzyuczelnianego wydziału Biotechnologii Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego przeprowadził jakościową analizę składu błon przetrwalnika *B. subtilis*. Analizę składu błony przeprowadzono z zastosowaniem metody TLC (ang. *thin layer chromatography*) ekstraktów z błon. Identyfikacja poszczególnych składników błon odbyła się wg. wzorców wykorzystanych przy TLC. Była to analiza głównie jakościowa, pozwalająca stwierdzić, jak rozkładają się poszczególne składowe błony — których elementów jest mniej, a których więcej bez wyrażenia tego w postaci liczbowej.

Na podstawie tych wyników oraz wyników opisanych w punkcie Model błon biologicznych *B. subtilis* zbudowałam model błony fosfolipidowej, który został w dalszych etapach wykorzystany jako model wewnętrznej błony przetrwalnika.

5.8.2. Orientacja białek receptora w błonie

Orientacje podjednostek — zarówno monomerów jak i dimerów — a także kompleksów części błonowych receptora GerA, obliczyłam za pomocą wspomnianego już narzędzia QMbrane (Studer, Biasini i Schwede, 2014) oraz serwera PPM 2.0 (ang. *Positioning of proteins in membranes*). Serwer ten oblicza, na podstawie struktury białka, jego rotacyjne i translacyjne pozycje w błonie (Lomize, Pogozheva, Joo, Mosberg i Lomize, 2012).

QMbrane, oprócz oceny jakości modelu, oblicza także szerokość segmentu błonowego oraz pseudoenergię potrzebną do transferu białka do błony. Na podstawie wartości pseudoenergii badane białko jest klasyfikowane jako białko błonowe, peryferyjne lub niebłonowe. Wyniki obliczeń serwera PPM obejmują szerokość błony, kąt pod jakim białko jest zorientowane w błonie, energię zanurzenia białka w błonie oraz segmenty transbłonowe i reszty aminokwasowe zanurzone w błonie dla obliczonej orientacji. Podczas obliczeń orientacji w serwerze PPM należy zaznaczyć, czy końce białek znajdują się wewnątrz cytoplazmy, czy na zewnątrz. Opcja ta dotyczy obu (N- oraz C-końca) końców i

nie można rozdzielić wyboru, jeżeli istnieje podejrzenie, że jeden koniec znajduje się wewnątrz, a drugi na zewnątrz.

5.8.3. Charmm-GUI

Charmm-GUI (Wu i inni, 2014) został przeze mnie wykorzystany do zbudowania modelu błony i zanurzenia w niej białka GerAB. Jest to serwer, który w swojej ofercie zawiera opcje budowania dwuwarstwy lipidowej oraz wygenerowania plików wsadowych do GROMACSa i innych programów obliczeniowych.

Model błony imitującej wewnętrzną błonę przetrwalnika zbudowałam na podstawie połączenia wyników otrzymanych w laboratorium prof. Banasia oraz informacji zawartych w opisanych w punkcie 3.9.2 wyników pracy badawczej, na podstawie których autorzy wykonali model błon biologicznych *B. subtilis*. Wyniki badań prof. Banasia zostały opisane w rozdziale 6.6.4.

W bibliotece lipidów Charmm-GUI brakowało niektórych długości łańcuchów kwasów tłuszczowych wymienionych w wynikach badań nad składem błon nieopłaszczonych przetrwalników, dlatego w przygotowaniu modelu błony lipidowej starałam się wybrać najbardziej podobne kwasy tłuszczowe, biorąc pod uwagę ich procentowy rozkład (tabela 14) wynikający z badań metodą TLC oraz proponowany przez (Nickels i inni, 2017) rozkład grup głównych. Ostateczny lipidowy model błony obliczony został przez algorytmy Charmm-GUI, na podstawie danych z wyżej wspomnianych eksperymentów i jest opisany w tabeli 2.

Kompleks receptora GerA zbudowany dimerów oraz kompleks zbudowany z monomerów zanurzyłam w modelu błony. Oba kompleksy GerA, ze względu na swoje nieregularne kształty zostały zanurzone metodą *replacement method*. Optymalizacja obu systemów odbyła się w temperaturze 303,15°K. Kompleks receptora GerA składający się z dimerów zorientowałam w błonie za pomocą wektora zbudowanego z dwóch reszt aminokwasowych: Leu116 łańcucha B oraz Gly260 łańcucha D. Następnie kompleks został przesunięty wzdłuż osi Z o 0.1272388995Å oraz obrócony względem osi X o 35°. Wartości translacji wynikają z obliczeń orientacji błony wykonanych w QMbrane oraz kilku prób „ręcznego” dopasowania orientacji błony do kompleksów. Dla kompleksu zbudowanego z monomerów punktami budującymi wektor były reszty aminokwasowe Pro266 łańcucha A i Ile159 łańcucha B. Przesunięcia i rotacje względem osi układu współrzędnych były takie

same jak w przypadku kompleksu dimerów. Rozmiar błony dla kompleksu w obu przypadkach wynosił 250x250 Å.

Tabela 2: Skład lipidowy modelu wewnętrznej błony, przeliczone przez Charmm-GUI dla błony o rozmiarach 250x250 Å.

Typ lipidu	Ilość w zewnętrznej warstwie błony	Ilość w wewnętrznej warstwie błony
DMPE	117	87
SOPE	39	29
YOPE	39	29
DOPE	78	58
TYCL1	78	116
TYCL2	78	116
TOCL1	78	58
TOCL2	78	58
DIGL	39	29

6. WYNIKI

6.1. Modele podjednostek receptora GerA

6.1.1. Podjednostka A

Białko GerAA, czyli podjednostka A receptora GerA, jest największym z białek tego receptora. Białko to zbudowane jest jednocześnie z integralnej domeny błonowej oraz domeny pozabłonowej. Praca nad teoretycznym modelem tego białka była niezwykle utrudniona, nie tylko ze względu na błonowy charakter podjednostki, ale też z powodu braku homologii tego białka do innych białek o znanej strukturze. GerAA wykazuje wysoką homologię do podjednostek A innych receptorów kiełkowania — przy czym, żaden z tych receptorów nie miał (aż do 2019 roku) opisanej struktury przestrzennej. Taki stan rzeczy znacząco utrudniał modelowanie homologiczne. Dodatkowo, wielkość białka (482 reszty aminokwasowe) dyskwalifikuje metody z rodzaju modelowania *ab initio*, które przy braku homologa wydawałyby się rozsądnym podejściem.

W związku z powyżej opisanymi trudnościami, podczas swojej pracy wielokrotnie poprawiałam model GerAA — każda nowo zdobyta informacja strukturalna była na bieżąco wykorzystywana do poprawy istniejącego modelu. Dzięki temu powstało kilka modeli, które dostarczyły danych dotyczących struktury oraz potencjalnych miejsc interakcji z germinantami. W tym podrozdziale chciałabym zaprezentować dwa najlepsze modele: pierwszy z nich został zbudowany przed opublikowaniem struktury domeny rozpuszczalnej jednego z homologów, drugi natomiast uwzględnia już te, niezwykle cenne, informacje.

6.1.1.1. Pierwszy model GerAA

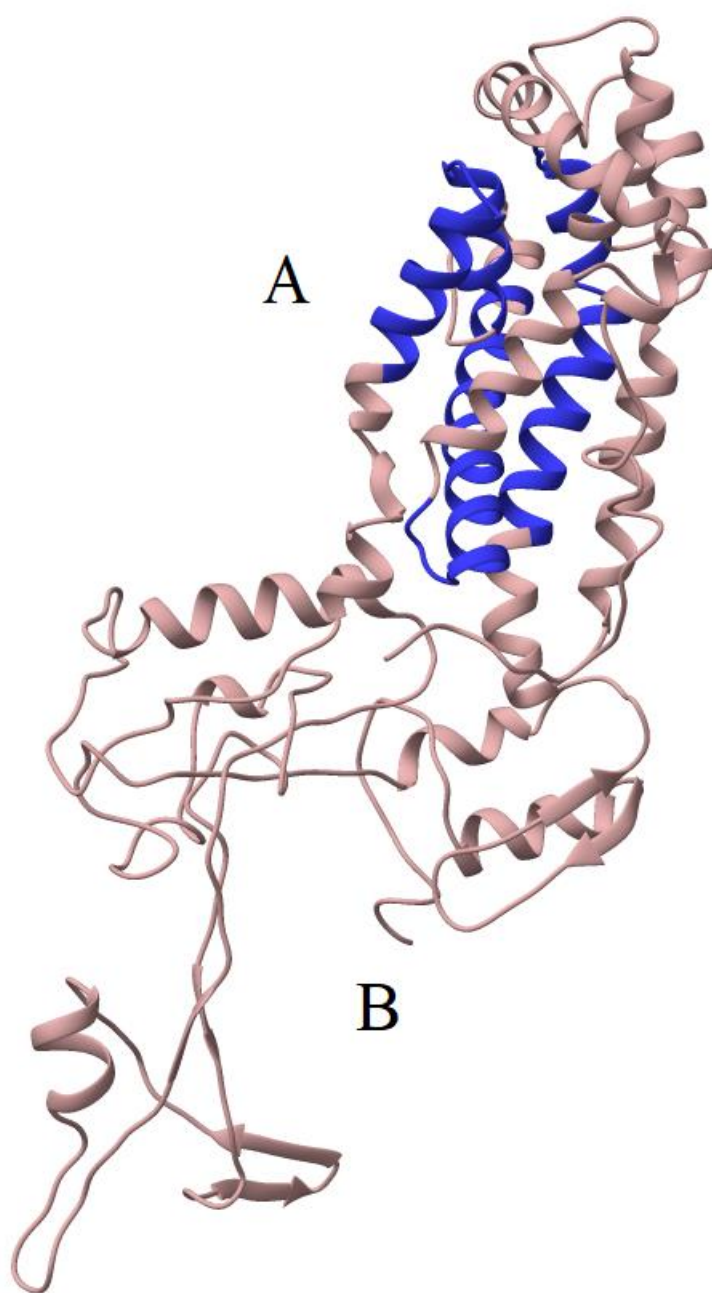
Pierwszy model GerAA utworzyłam z wykorzystaniem wieloszablonowego modelowania homologicznego, a jako szablony użyłam trzech błonowych transporterów bakteryjnych, które zostały opisane w rozdziale 5.4 *Przewidywanie struktur trzeciorzędowych*. Modelowanie homologiczne na wielu szablonach zostało wykorzystane ze względu na brak, w momencie budowania pierwszych modeli, struktur homologicznych, które można byłoby wykorzystać w modelowaniu.

Użycie serwera I-Tasser pozwoliło mi na wytypowanie białek wystarczająco podobnych — zarówno pod względem struktury jak i funkcji — by mogły służyć jako szablony dla modelu GerAA. Na podstawie wyników wyszukiwania szablonów wytypowałam trzy białka: 3K07, 4MT1 oraz 4DNT. Są to duże, znacznie większe od GerAA, białka błonowe transporterów bakteryjnych.

Wyniki modelowania homologicznego GerAA na podstawie całych białek użytych jako szablon były wynikami o średniej lub niskiej jakości, co prawdopodobnie wynikało z rozmiaru szablonów. W związku z tym przeprowadziłam lokalne zestawienia sekwencji dla podjednostki A i każdego z szablonów. Pozwoliło to na wytypowanie fragmentów najbardziej podobnych do białka docelowego. Fragmenty szablonów zostały zebrane i poddane wielosekwencyjnemu zestawieniu wraz z sekwencją białka GerAA w algorytmach MUSCLE i MAFFT. Wyniki zestawienia wykonanego przez algorytm MAFFT dawały lepsze wyniki, dlatego zostały przeze mnie wykorzystane w kolejnych krokach modelowania struktury. Identyczność sekwencji pomiędzy szablonami i białkiem docelowym uzyskana z tego zestawienia wynosiła 19,295%.

Tak dobrane fragmenty szablonów zostały wykorzystane w modelowaniu homologicznym struktury białka GerAA, dając model o wartości GA341 równej 0,90888, co jest wysokim wynikiem, biorąc pod uwagę, że im bliżej jedności tym lepsza jest jakość uzyskanego modelu. Wartości RMSD dla tego modelu w porównaniu do szablonów wynoszą kolejno: 1,38Å dla 3K07, 0,93Å dla 4MT1 oraz 0,72Å dla 4DNT.

Struktura pierwszego modelu GerAA została przedstawiona na rysunku 5. Model wyraźnie dzieli się na dwie części: część ustrukturyzowaną, zbudowaną z 6 dużych helis, która odpowiada domenie transbłonowej oraz część nieustrukturyzowaną — N-domenę pozabłonową. Nieuporządkowanie fragmentu N-terminalnego białka prawdopodobnie wynika z doboru szablonów — szablony, użyte do modelowania homologicznego, to transportery błonowe zbudowane głównie z helis transbłonowych. N-koniec białka GerAA jest domeną niebłonową.



Rysunek 5: Pierwszy model GerAA wraz z zaznaczonymi fragmentami transbłonowymi (kolor niebieski). Przewidywana domena błonowa została oznaczona literą A. Na modelu widać dużą liczbę nieuporządkowanych pętli w domenie pozabłonowej białka (oznaczone jako fragment B).

6.1.1.2. Drugi model strukturalny GerAA – modelowanie homologiczne

Nieocenione w pracach nad poprawą jakości modelu GerAA okazało się opublikowanie w 2019 roku struktury fragmentu podjednostki A receptora GerK z *Bacillus megaterium* (Li i inni, 2019). Struktura domeny pozabłonowej podjednostki A została rozwiązana z wykorzystaniem anomalnej dyfrakcji pojedynczej fali (ang. *single-wavelength anomalous dispersion* SAD). Ze względu na wysoką homologię do podjednostki A receptora GerA, struktura ta została wykorzystana przeze mnie w modelowaniu homologicznym fragmentu niebłonowego podjednostki A receptora GerA.

Ponieważ dostępny był jedynie fragment niebłonowy homologa, podczas prac nad przewidywaniem *drugiego* modelu struktury białka GerAA przyjęłam następującą strategię: białko docelowe podzieliłam na dwie części — N-terminalną domenę rozpuszczalną (reszty 1-256) oraz domenę błonową (reszty 238-482). Zakres domeny rozpuszczalnej zdefiniowałam na podstawie podobieństwa do homologa 6O59, natomiast integralna domena błonowa została zidentyfikowana na podstawie wyników przewidywań regionów transmembranowych białka (tabela 3). Warto zauważyć, że oba regiony nachodzą na siebie – od reszty 238 do reszty 256. Jest to celowy zabieg, który ułatwił połączenie obu regionów w kolejnych krokach przewidywania struktury.

Jak już wspomniałam, białko opisywane kodem 6O59 zostało wybrane przeze mnie jako szablon podczas modelowania struktury domeny rozpuszczalnej podjednostki A. Na podstawie zestawień sekwencji do modelowania wybrałam łańcuch A homologa. Modelowanie homologiczne (wykonane za pomocą oprogramowania Modeller opisanego w rozdziale 5.4 *Przewidywanie struktur trzeciorzędowych*) tej części białka dało bardzo wiarygodne wyniki. Wartość RMSD pomiędzy domeną rozpuszczalną a łańcuchem A białka 6O59 wynosi 0,55Å. Podobieństwo sekwencji do białka wzorcowego wynosi 50%, a identyczność sekwencji docelowej i szablonu jest na poziomie 36,8%.

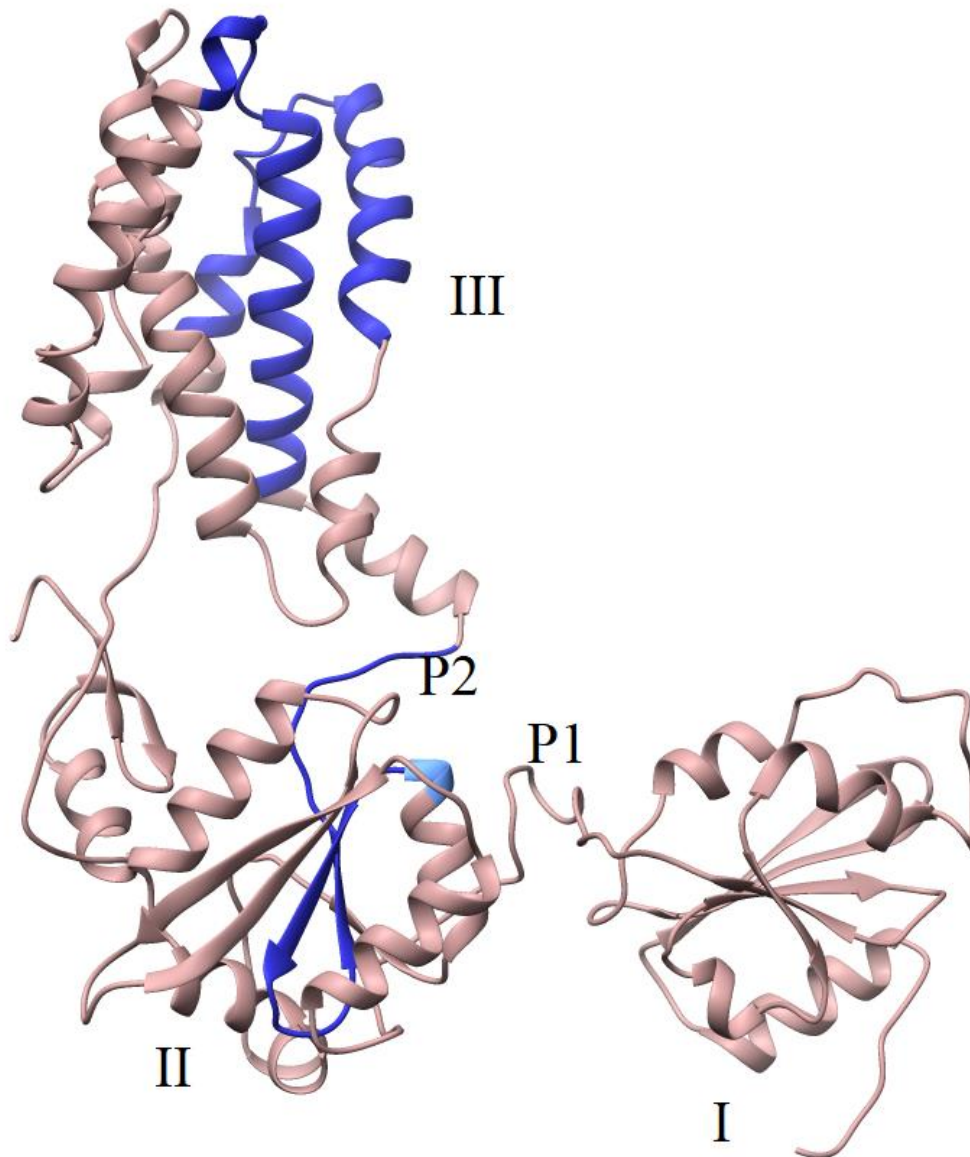
W kolejnym kroku pracy nad modelem przygotowałam model integralnej części błonowej białka GerAA. Została ona wymodelowana przy użyciu białka 3WDO jako szablonu (wybór szablonu został wyjaśniony w rozdziale Materiały i Metody), w wyniku czego otrzymałam model o wysokiej jakości (rysunek 6). Globalne zestawienie sekwencji szablonu oraz białka docelowego wykazało 11,7% identyczność i 20,0% podobieństwa z białkiem wzorcowym, dlatego zdecydowałam się na przeprowadzenie lokalnego zestawienia PSA. Lokalne PSA wykazało 23,1% identyczność i 38,5% podobieństwo z regionem 3WDO składającym się z reszt 5-204. Na podstawie tych wyników do

modelowania homologicznego wykorzystałam lokalne zestawienie PSA oraz sugerowany tym zestawianiem fragment struktury szablonu. Wartość RMSD modelu nałożonego na szablon użyty podczas modelowania, którym był region 5-205 białka 3WDO, wynosi 0,57Å.

Obie części modelu następnie połączyłam za pomocą narzędzia *alignment tool* dostępnego w programie PyMol, a następnie przeprowadziłam modelowanie całego już białka, używając połączonego modelu jako szablonu. Modelowanie oraz optymalizacja modelu zostały przeprowadzone przy wykorzystaniu programu Modeller.

Ponieważ białko GerAA jest w połowie integralnym białkiem błonowym, ocenę jakości modelu przeprowadziłam przy użyciu algorytmu Qmbrane, narzędzia stworzonego do oceny białek błonowych. Wynik tej ewaluacji kwalifikuje model jako model dobrej jakości i wysokiej wiarygodności, z następującymi wynikami: *normal score* = 0,54, *Z-score* = -1,58 oraz 91,88% reszt opisanych jako korzystne na wykresie Ramachandrana. Wyniki oceny potwierdzają również, że białko GerAA jest białkiem częściowo błonowym.

W porównaniu z pierwszym modelem GerAA, który został przeze mnie opisany w podrozdziale 6.1.1.2, widać, że model uzyskany z wykorzystaniem szablonów 3WDO oraz 6O59 również charakteryzuje się dwiema domenami, przy czym domena niebłonowa zbudowana jest z wyraźnie ustrukturyzowanych fragmentów. Wszystkie badania przeprowadzone do tej pory na modelu pierwszym (dokowania germinantów, badania dotyczące dimerów i trimerów podjednostek, badania ruchliwości białek, badania nad orientacją białek receptora w błonie) zostały przeze mnie powtórzone dla nowego (drugiego) modelu GerAA. W dalszej części rozdziału będę wyraźnie zaznaczać, którego modelu dotyczą omawiane wyniki.



Rysunek 6: Model białka GerAA (tzw. *drugi model*). Domena III jest integralną domeną błonową. Kolorem niebieskim zaznaczone zostały przewidywane fragmenty błonowe. Wyraźnie widać, że fragment struktury przewidziany jako pierwsza helisa błonowa jest częścią domeny II i nie przyjmuje struktury α -helisy. Domeny I oraz II wchodzi w skład dużej domeny pozabłonowej i zbudowane są z helis oraz β -kartek. Pętla P1 łączy ze sobą domenę I oraz domenę II. Pętla P2 łączy ze sobą części błonową oraz pozabłonową białka.

6.1.1.3. Opis struktury modelu GerAA – na podstawie modelu drugiego

Model GerAA wykazuje wyraźne zróżnicowanie na integralną domenę błonową i domenę pozabłonową — domeny te są ze sobą połączone łącznikiem o długości 10 reszt aminokwasowych (reszty 256-266). Interesującym jest, że część tego linkera została przewidziana jako helisa TM podczas przewidywania regionów błonowych (tabela 3). Domena transbłonowa białka umiejscowiona jest pomiędzy dużą domeną N-terminalną (domena pozabłonowa) a C-terminalnym ogonem zbudowanym z dwóch β -katek i jednej α -helisy (rysunek 6). Wyniki przewidywań regionów transbłonowych wskazują na to, że oba końce białka będą znajdować się po wewnętrznej stronie błony wewnętrznej przetrwalnika.

Domena błonowa białka obejmuje wysoce konserwowany motyw PFPP (reszty 324-327), który, na podstawie przewidywań regionów błonowych, znajduje się pomiędzy dwiema transmembranowymi helisami. Przewidywanie struktur drugorzędowych i trzeciorzędowych pokazuje, że V323 i P324 przyjmują strukturę pętli łączącej dwie helisy, natomiast reszty FPP z motywu PFPP stanowią część tej drugiej helisy łączonej przez wspomniany linker zbudowany z reszt V323 i P324. Długość helis oraz łącznika różni się pomiędzy wynikami przewidywań struktur drugorzędowych oraz uzyskanym przeze mnie modelem struktury trzeciorzędowej (rysunek 8).

Przewidywania dotyczące regionu transbłonowego, przedstawione w tabeli 3, wskazują, że fragment błonowy białka składa się z czterech długich helis TM. Modele drugorzędowe i trzeciorzędowe pokazują natomiast, że domena ta zbudowana jest nawet z dziesięciu helis błonowych o różnej długości. Niektóre z nich są bardzo krótkie, inne — długie. Różnice te tłumaczyć można niedokładnością szacowania występowania regionów transbłonowych oraz niedokładnościami związanymi z przewidywaniami struktur — jakością szablonu użytego w modelowaniu homologicznym oraz niedoskonałością algorytmów służących do przewidywań struktur drugorzędowych. W literaturze dotyczącej hydrofobowości GerAA uznaje się, że białko zawiera 4-6 helis transbłonowych.

W pozabłonowej części modelu GerAA, podobnie jak w użytym do modelowania szablonie, można wyróżnić dwie domeny o niemal identycznej strukturze. Pierwsza domena składa się z czterech krótkich helis i czterech katek beta, natomiast druga — z pięciu helis i czterech β -katek. Struktury alfa i beta każdej z domen przeplatają się, tworząc strukturę

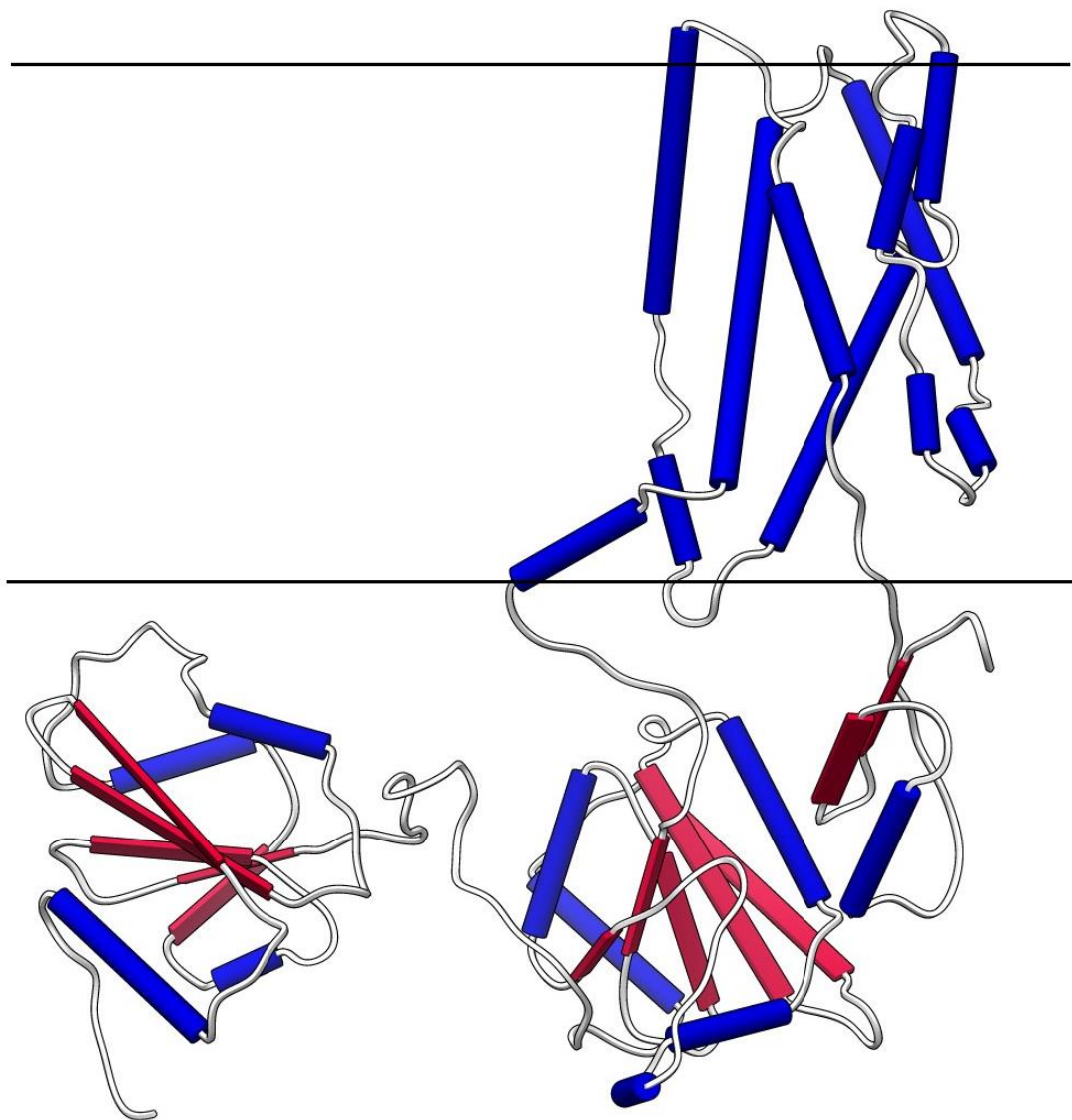
nazywaną $\alpha\beta$ -kanapką (ang. *$\alpha\beta$ -sandwich*) — struktury drugorzędowe pokazane są na rysunku 7.

Porównanie struktur trzeciorzędowych i drugorzędowych GerAA przedstawiłam na rysunku 8. Porównanie to wykazuje dużą zbieżność przewidywań w domenie N-terminalnej i niewielkie rozbieżności w domenie hydrofobowej, gdzie duże α -helisy (w tym helisy błonowe) są podzielone na kilka krótszych struktur w modelu struktury trzeciorzędowej. Dodatkowo, przewidywany ogon na C-końcu białka, w modelu struktury trzeciorzędowej przyjmuje postać dwóch krótkich β -kartek. To porównanie, wraz z wynikami oceny jakości struktury, wskazuje na wysoką dokładność oraz jakość trójwymiarowego modelu struktury. Na tej podstawie model ten można wykorzystać do dalszych badań nad receptorem GerA.

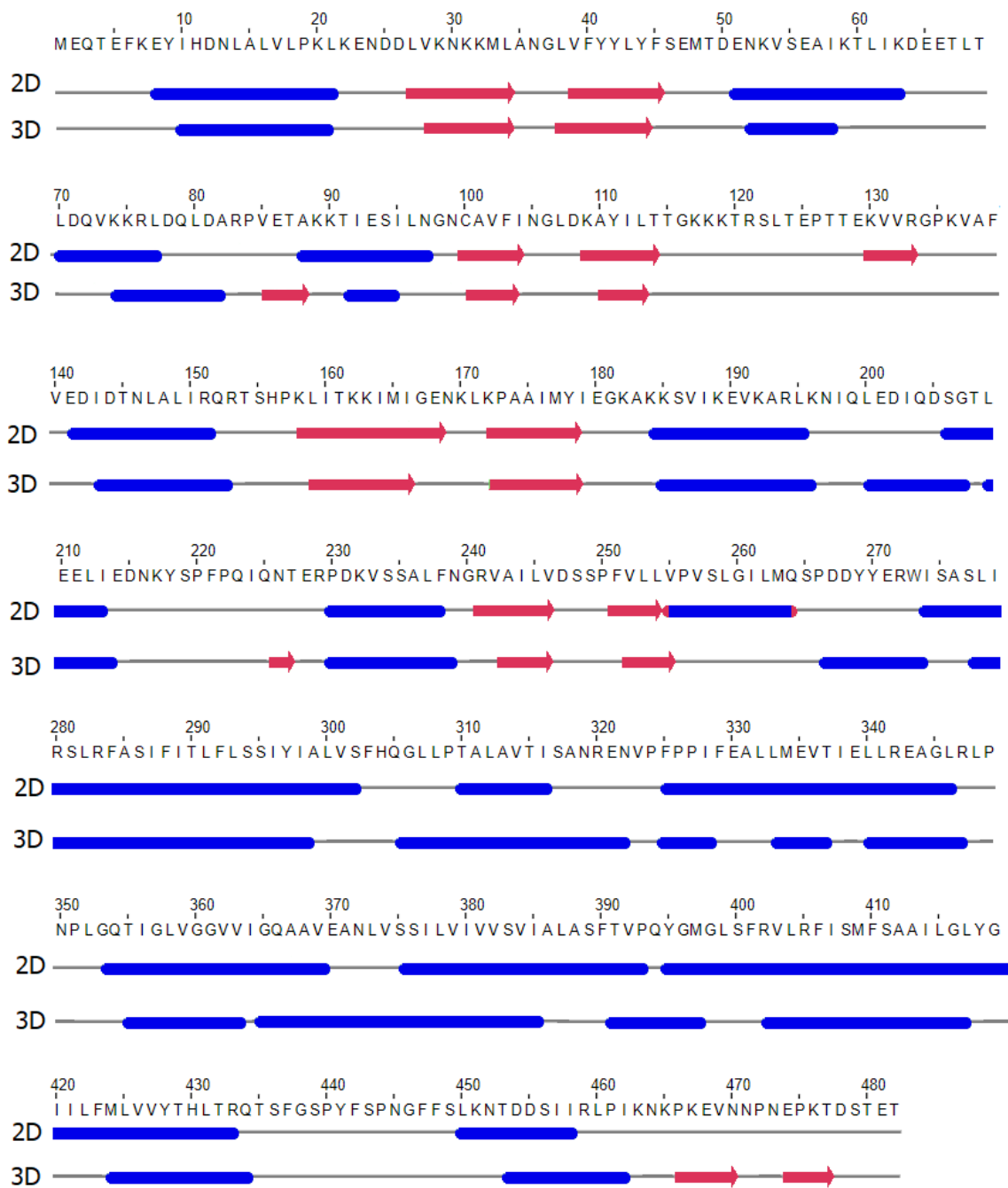
Tabela 3: Regiony transmembranowe GerAA przewidywane za pomocą trzech algorytmów: TMap, TMPred i TMHMM. Wykorzystane metody przewidywań regionów transbłonowych wskazują cztery helisy transbłonowe w białku GerAA.

GerAA	TMAP	TMPred	TMHMM
Helisa 1	238-246	242-265	242-264
Helisa 2	280-308	284-300	279-301
Helisa 3	360-388	373-392	374-396
Helisa 4	402-430	407-429	406-428

W tabeli 3 umieściłam wyniki sugerujące, że region białka GerAA opisany jako helisa 1 przenika przez błonę i rozciąga się od mniej więcej reszty 238 do reszty 265, w zależności od tego, który algorytm bierzemy pod uwagę. O ile w strukturze trzeciorzędowej *pierwszego modelu GerAA* możemy zauważyć, że ten region rzeczywiście przyjmuje strukturę helisy, o tyle przewidywania struktur drugorzędowych oraz *drugi model* temu przeczą. Na rysunku 6 widać, że ta część modelu znajduje się w domenie niebłonowej, co więcej dominują w niej struktury β . Przewidywania struktury drugorzędowej (rysunek 8) z kolei pokazują, że te reszty aminokwasowe budują dwie β -kartki, po których następuje α -helisa. Rozbieżności te mogą wynikać z niedokładności algorytmów przewidywania regionów transbłonowych. Biorąc pod uwagę fakt, że ta część modelu białka została zbudowana na podstawie podobieństwa do bliskiego homologa, a przewidywania struktur drugorzędowych dość blisko pokrywają się z modelem, sensownym wydaje się przyjęcie założenia, że: i) hipotetyczna „helisa 1” nie jest w rzeczywistości helisą, ii) być może nie jest to fragment błonowy białka.



Rysunek 7: Model GerAA przedstawiający struktury drugorzędowe w domenie błonowej i domenie pozabłonowej. Helisy zaznaczone są kolorem niebieskim, β -kartki kolorem czerwonym. Czarne poziome linie reprezentują, bardzo poglądowo, wewnętrzną błonę przetrwalnika.



Rysunek 8: Porównanie przewidywań struktury drugorzędowej(2D) oraz trzeciorzędowej(3D) dla białka GerAA. Kolorem niebieskim zaznaczone są α -helisy, natomiast czerwonym - β -kartki.

6.1.2. Model podjednostki B

Podjednostka B receptora GerA jest integralnym białkiem błonowym, znajdującym się w wewnętrznej błonie przetrwalnika. Błonowy charakter białka GerAB w połączeniu z niedostępnością skutecznego przeciwciała anty-GerAB sprawiają, że ustalenie struktury krystalicznej tej podjednostki za pomocą znanych metod eksperymentalnych jest niezwykle trudne. W przypadku, gdy informacje strukturalne dotyczące białka są bardzo ubogie, ale niezwykle istotne w badaniach nad jego funkcją i zachowaniem, najlepszym rozwiązaniem jest sięgnięcie po dostępne metody budowy modeli strukturalnych *in silico*. Do tej pory w literaturze nie pojawiły się wzmianki o strukturze podjednostek B inne niż modele teoretyczne bazujące na nikłej homologii, jaką te białka wykazują do rodziny transporterów bakteryjnych APC.

W ostatnim czasie w literaturze pojawiły się wyniki dwóch autonomicznych badań teoretycznych skupionych na problemie struktury białka GerAB, jego interakcji z błoną komórkową oraz z germinantami (Blinker, Vreede, Setlow i Brul, 2021; Artzi i inni, 2021). Badania te były prowadzone niezależnie od badań opisywanych w niniejszej rozprawie, dają jednak bardzo zbliżone rezultaty do prezentowanych przeze mnie wyników, co zostało opisane w rozdziale 7 *Dyskusja*.

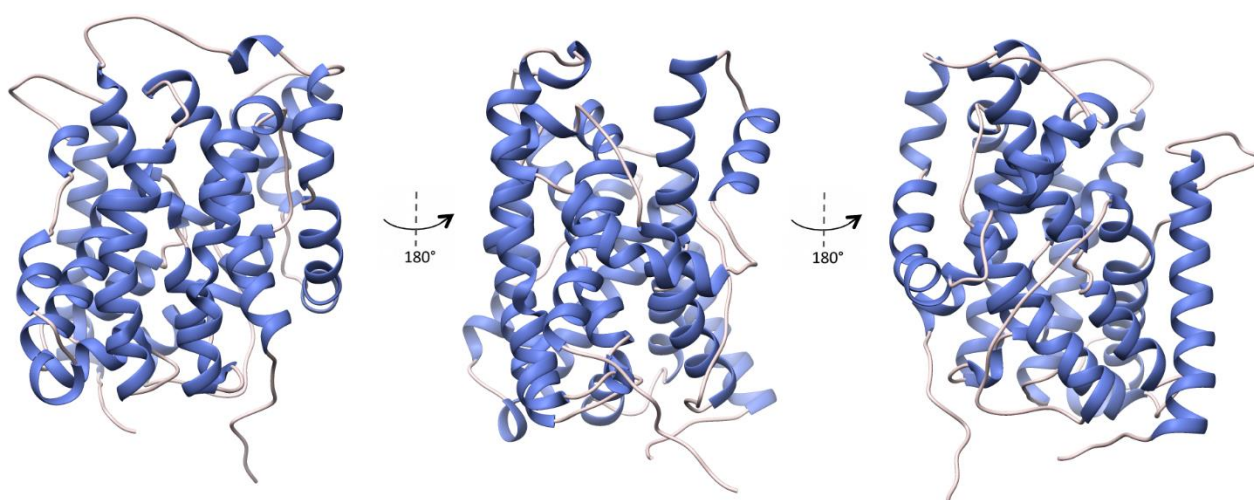
6.1.2.1. Modelowanie struktury GerAB

GerAB jest białkiem zbudowanym z 365 reszt aminokwasowych i podobnie jak pozostałe podjednostki receptora GerA, wykazuje bliską homologię jedynie do innych białek B receptorów kiełkowania. Homologi te, podobnie jak GerAB nie mają znanej struktury krystalicznej, z której można byłoby skorzystać podczas metod modelowania homologicznego. W tym przypadku jednak, można wziąć pod uwagę nikłą homologię podjednostki B do nadrodziny transporterów bakteryjnych APC.

Z tego powodu, w celu zbudowania modelu teoretycznego struktury GerAB, przeprowadziłam wieloszablonowe modelowanie homologiczne w oparciu o szablony znalezione za pomocą algorytmów I-Tasser o identyczności sekwencji oscylującej wokół 19%. Szablony — białka z rodziny APC — zostały opisane w rozdziale 5 *Materiały i metody*.

Ocena modelu GerAB dokonana za pomocą narzędzia Qmbrane pokazuje, że jest to wiarygodny model o wysokiej jakości. Algorytm oceny wskazuje następujące wyniki:

normal score = 0,62 i *Z-score* = -0,72. Liczba korzystnych reszt na wykresie Ramachandrana plasuje się na poziomie 89,26%, co jest wysoką wartością, biorąc pod uwagę fakt, że mamy do czynienia z integralnym białkiem błonowym. Białka błonowe w większości wypadków są gorzej oceniane na wykresie Ramachandrana, modelowanie jest trudniejsze i znane algorytmy oceniania często dają takim białkom słabsze oceny. Wartości RMSD modelu GerAB nałożonego na każdy z szablonów są następujące: 4,0Å do 3L1L, 2,7Å do 3GIA, 2,6Å do 3GI8 oraz 2,6Å do 3GI9. Patrząc na porównanie przewidywań modeli drugo- i trzeciorzędowych (rysunek 10) oraz ocenę modelu struktury trzeciorzędowej przyjął, że przewidywane modele struktury 3D są wystarczająco dokładne, by przeprowadzić za ich pomocą kolejne badania.



Rysunek 9: Model GerAB pokazany z różnych perspektyw.

6.1.2.2. Opis struktury modelu GerAB

Model strukturalny podjednostki B receptora GerA zbudowany jest z dziewiętnastu helis (rysunek 11), podczas gdy przewidywania struktury drugorzędowej wskazują, że białko to składa się z czternastu helis. Białko GerAB przyjmuje charakterystyczną dla białek błonowych budowę opartą na licznych α -helisach. Porównanie przewidywań struktur trzeciorzędowych i drugorzędowych pokazuje, że niektóre z większych helis przewidzianych przez algorytmy do przewidywania struktur drugorzędowych, wydają się być rozbite na mniejsze w zbudowanym przez mnie trójwymiarowym modelu GerAB. Może to wynikać z niskiego poziomu homologii pomiędzy białkiem docelowym a jego

wzorcami. Porównanie struktur trzeciorzędowych i drugorzędowych GerAB pokazane jest na rysunku 10.

Przewidywania struktur drugorzędowych, oraz jeden z modeli opublikowanych niedawno przez Blinker i in. (Blinker, Vreede, Setlow i Brul, 2021), sugerują, że w białku B występuje bardzo krótka β -karkta (reszty 175-177). W modelu prezentowanym w tej pracy nie występuje struktura beta — podobnie w innych modelach prezentowanych we wspomnianej publikacji. Porównując tę β -karkkę do przewidywań regionów transbłonowych widzimy, że jest to fragment pozabłonowy, prawdopodobnie występujący po zewnętrznej stronie błony przetrwalnika. W modelu struktury trzeciorzędowej białka widzimy długą pętlę obejmującą reszty 266-279, według przewidywań struktur drugorzędowych w tym miejscu powinna znaleźć się długa helisa 267-302. W utworzonym przeze mnie modelu struktury trzeciorzędowej helisa ta jest krótsza i rozpoczyna się resztą 279. Dodatkowo helisa 84-107 obserwowana w strukturze drugorzędowej — jest to helisa transbłonowa, jak pokazuje tabela 4 — w modelu dzieli się na dwie mniejsze helisy połączone krótkim łącznikiem.

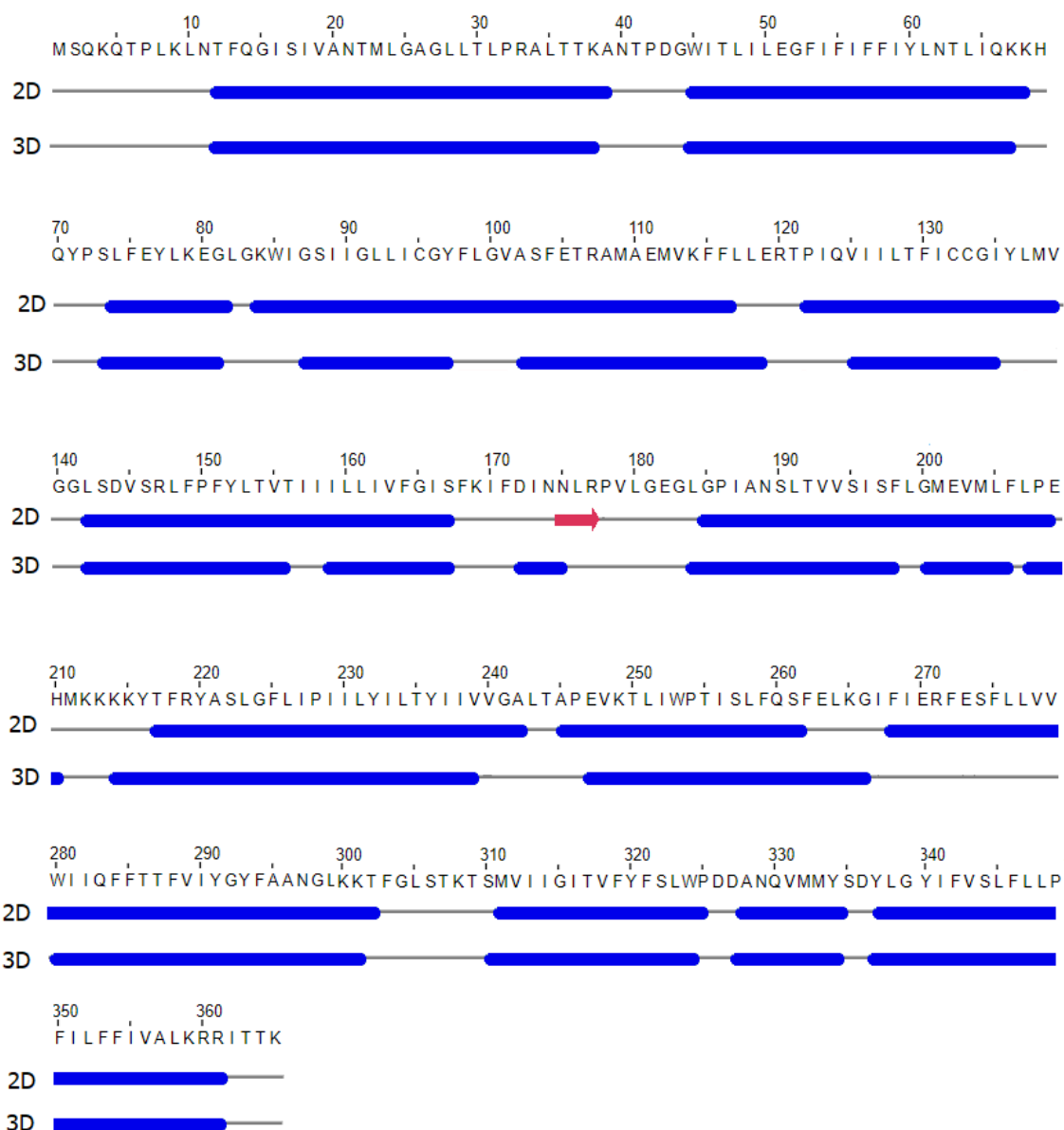
6.1.2.3. Przewidywanie regionów błonowych podjednostki B

Dane dostępne w literaturze podają, że oba końce białka GerAB znajdują się we wnętrzu przetrwalnika (Cooper i Moir, 2011). Helisy transbłonowe przewidywane przez dedykowane algorytmy zostały zebrane w tabeli 4. Podobnie jak w danych literaturowych, przewidywania te wskazują 10 helis perforujących błonę. Porównując te przewidywania ze strukturami drugo- i trzeciorzędowymi, można wyraźnie zauważyć, że helisy transmembranowe są połączone ze sobą hydrofilowymi regionami zbudowanym przez krótkie pętle i/lub helisy. Wspomniana wcześniej β -karkta przewidziana przez algorytm JPred znajduje się między helisami 5 oraz 6 we fragmencie białka, które znajduje się poza błoną od zewnętrznej strony błony.

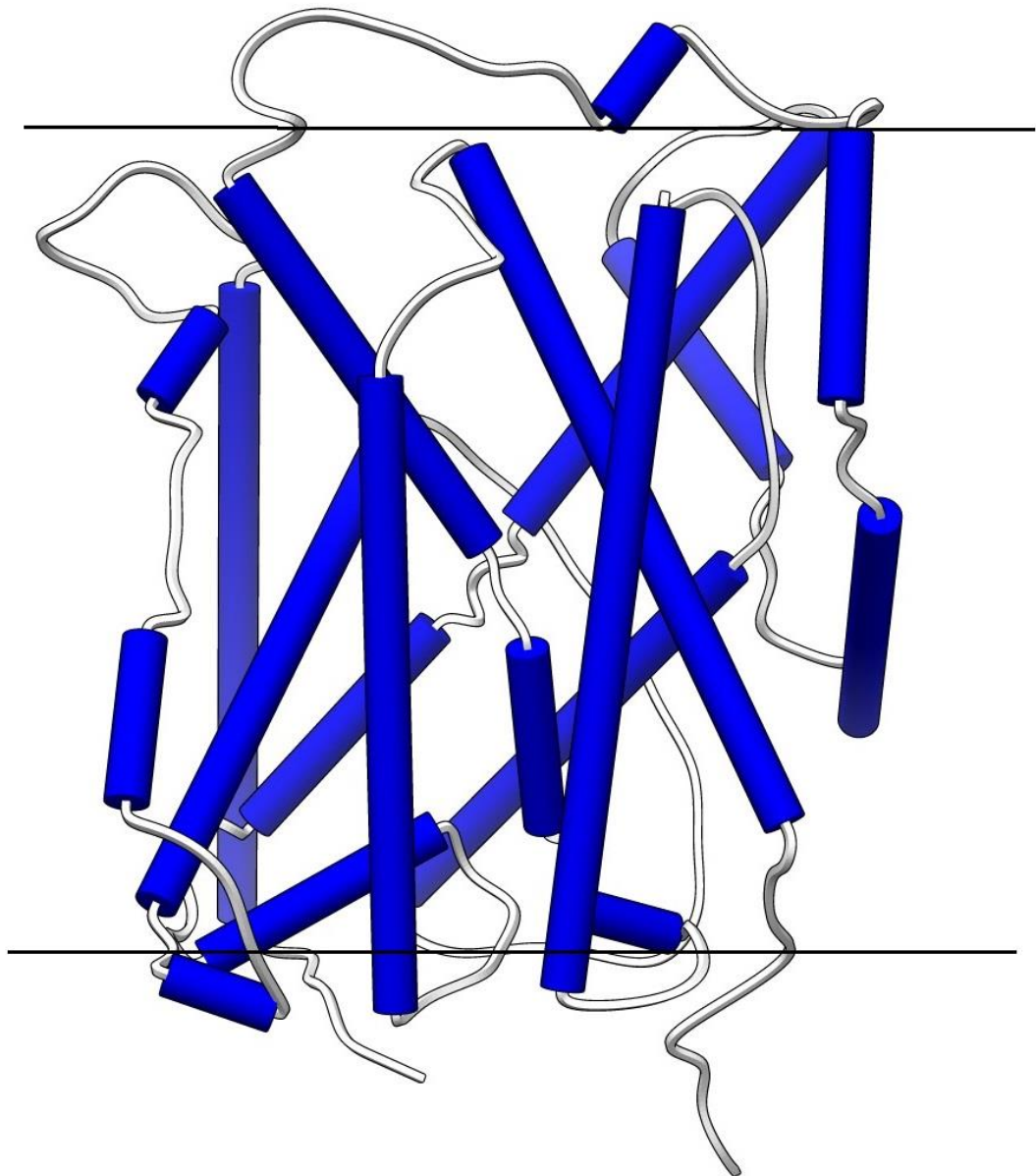
Tabela 4: Regiony transmembranowe GerAB przewidywane za pomocą trzech algorytmów: TMap, TMPred i TMHM. Wszystkie przewidywania dla GerAB potwierdzają, że białko to składa się z 10 helis błonowych.

GerAB	TMAP	TMPred	TMHMM
Helisa 1	13-37	13-31	13-35
Helisa 2	43-66	45-61	45-64
Helisa 3	87-107	86-104	85-106

Helisa 4	115-135	125-142	121-143
Helisa 5	146-174	148-168	148-170
Helisa 6	185-209	186-206	185-207
Helisa 7	221-249	223-239	220-242
Helisa 8	269-297	275-294	276-295
Helisa 9	306-326	303-325	302-324
Helisa 10	338-358	337-356	339-358



Rysunek 10: Porównanie przewidywań struktur drugorzędowych (2D) oraz trzeciorzędowych (3D) białka GerAB. Przewidywania struktur drugorzędowych zostały wykonane za pomocą algorytmów serwera JPred.



Rysunek 11: Model białka GerAB z zaznaczonymi strukturami drugorzędowymi. Helisy zaznaczone są kolorem niebieskim, natomiast czarne poziome linie reprezentują poglądowo umiejscowienie krańców wewnętrznej błony przetrwalnika.

6.1.3. Model podjednostki C

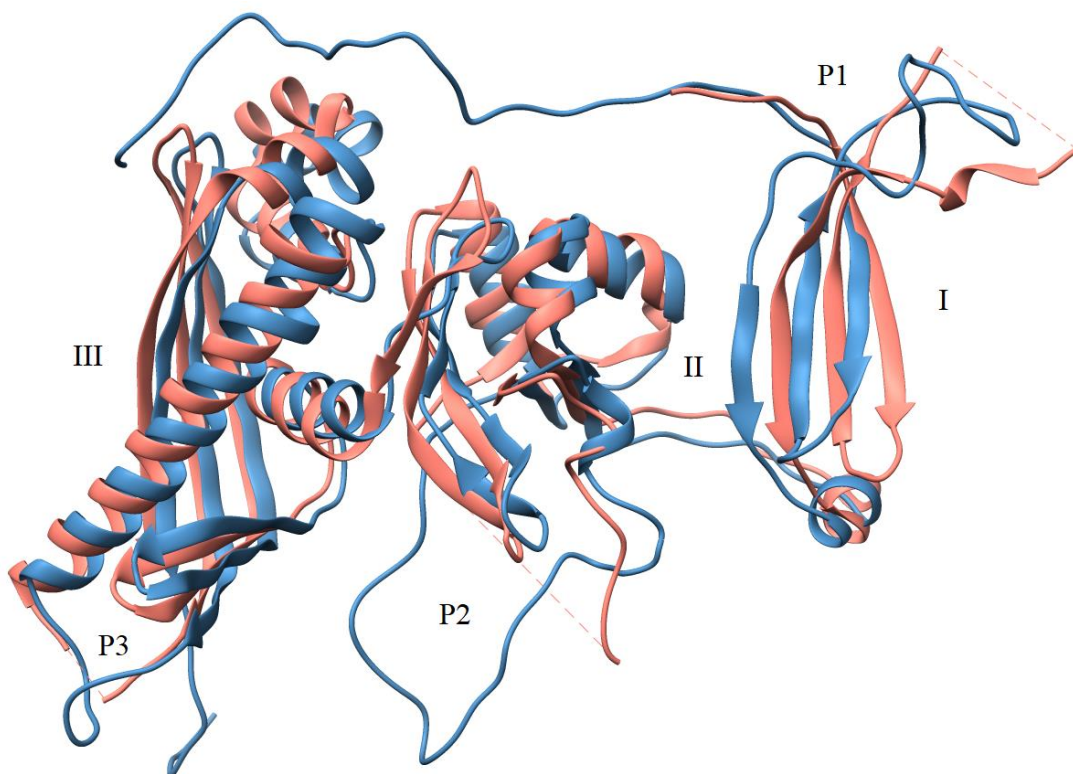
6.1.3.1. Modelowanie

Podjednostka C receptora kiełkowania GerA jest lipoproteiną i najprawdopodobniej jest połączona z błoną za pomocą kotwicy diacyloglicerolowej, a reszta cysteiny w pozycji 18 białka odpowiada za interakcje kotwicy z wewnętrzną błoną przetrwalnika. Białko GerAC wykazuje silną homologię strukturalną do białka podjednostki C receptora kiełkowania GerB *Bacillus subtilis* o znanej strukturze. Globalne zestawienie sekwencji wykazuje 33% identyczność pomiędzy sekwencjami białka docelowego i homologa.

Struktura homologa, której użyłam do przewidywania modelu struktury GerAC jest pozbawiona pierwszych 24 reszt aminokwasowych z N-końca. Są to najbardziej hydrofobowe reszty aminokwasowe tego białka — wśród nich cysteina 18. Brak tak dużego fragmentu białka użytego jako szablon w modelowaniu skutkuje występowaniem długiej, nieuporządkowanej pętli w modelu białka docelowego (rysunek 12). Jak podają badacze (Li, Setlow, Setlow i Hao, 2010), struktura krystaliczna białka GerBC wykazuje się specyficzną topologią, która wydaje się nie mieć homologa wśród innych znanych struktur białek.

Homolog podjednostki C w roztworze przyjmuje formę monomeru, natomiast w kryształach, gdzie upakowanie jest duże, zaobserwowano, że ulega on dimeryzacji. Strukturę dimeru homologa oraz potencjalnego dimeru białka GerAC szerzej opisałam w rozdziale poświęconym dimerom podjednostek (6.4.2.3 *Dimer GerAC*).

Model struktury trzeciorzędowej GerAC wykazuje duże podobieństwo do przewidywań struktury drugorzędowej, porównanie tych przewidywań przedstawiłam na rysunku 13. W modelu struktury trzeciorzędowej brakuje pierwszej helisy przewidzianej przez serwer JPred, β -kardki w miejscu 65-72 oraz helisy 151-164. Ponadto, kilka dłuższych struktur beta widocznych w strukturze drugorzędowej w modelu 3D jest podzielonych na dwie/trzy krótsze struktury. Wartość RMSD dla modelu nałożonego na szablon jest równa 0,5Å, co wskazuje na bardzo duże podobieństwo obu struktur oraz wysoką jakość tego modelu. Przypuszczenia te wydają się być potwierdzone przez narzędzia oceny QMEANDisCo, zgodnie z którymi *normal score* tego modelu jest równy 0,54, *Z-score* = -1,56, a *QMEANDisCo* = $0,55 \pm 0,5$. Reszty korzystne według wykresu Ramachandrana stanowią 93,26% reszt aminokwasowych tego białka.



Rysunek 12: Model białka GerAC (kolor niebieski) z nałożoną strukturą szablonu. Na ilustracji zaznaczone są trzy domeny – I, II, III – białka oraz trzy pętle (P1, P2 i P3) łączące te domeny. Wyraźnie widać, że w strukturze szablonu brakuje struktur nieuporządkowanych, czyli wspomnianych pętli. Dodatkowo widać nieobecność pierwszych reszt aminokwasowych odpowiedzialnych za przyłączenie białka do błony

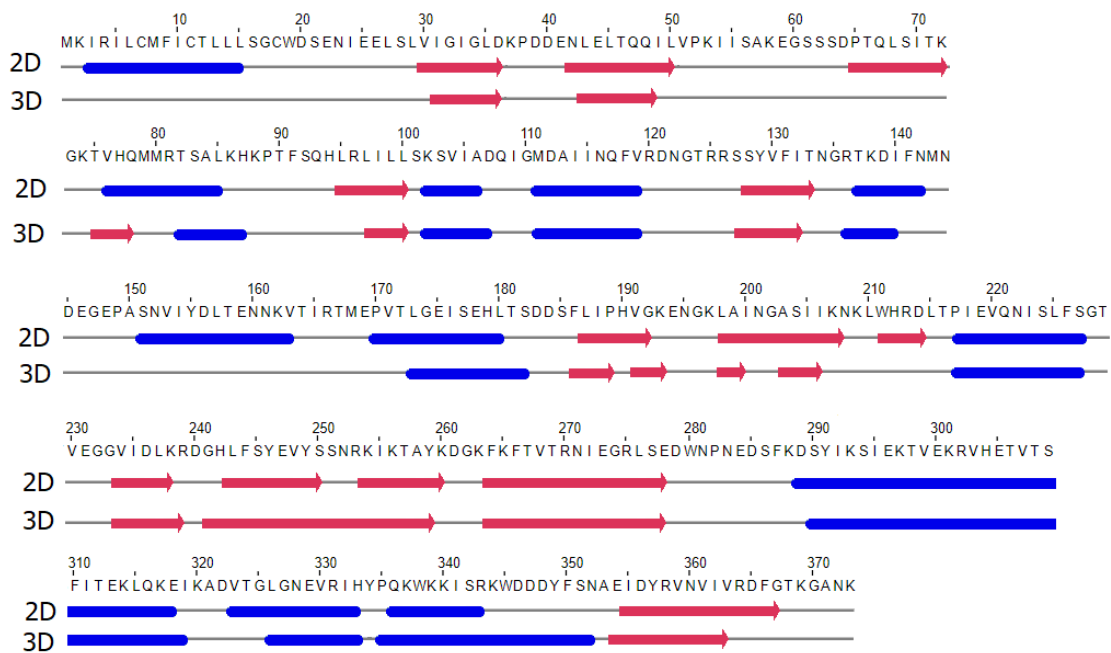
6.1.3.2. Opis struktury podjednostki C

Struktura białka GerAC, podobnie jak struktura jego homologa, wykazuje wyraźny podział na trzy domeny, które połączone są ze sobą nieuporządkowanymi regionami (rysunek 14). Pierwsza domena, domena N-końcowa, zbudowana jest z trzech równoległych β -krotek i jest zakończona krótką α -helisą (reszty 82-86). Druga, domena (umiejscowiona pomiędzy domenami I oraz III) składa się z sześciu struktur β i czterech krótkich α -helis. Pomiedzy czwartą i piątą helisą znajduje się długi nieuporządkowany region (P2), który rozciąga się od reszty 141 do reszty 172. Domeny I oraz II połączone są ze sobą krótką pętlą. Trzecia domena, C-terminalna, rozpoczyna się krótką helisą α , za którą znajdują się cztery β -karkki i trzy inne α -helisy.

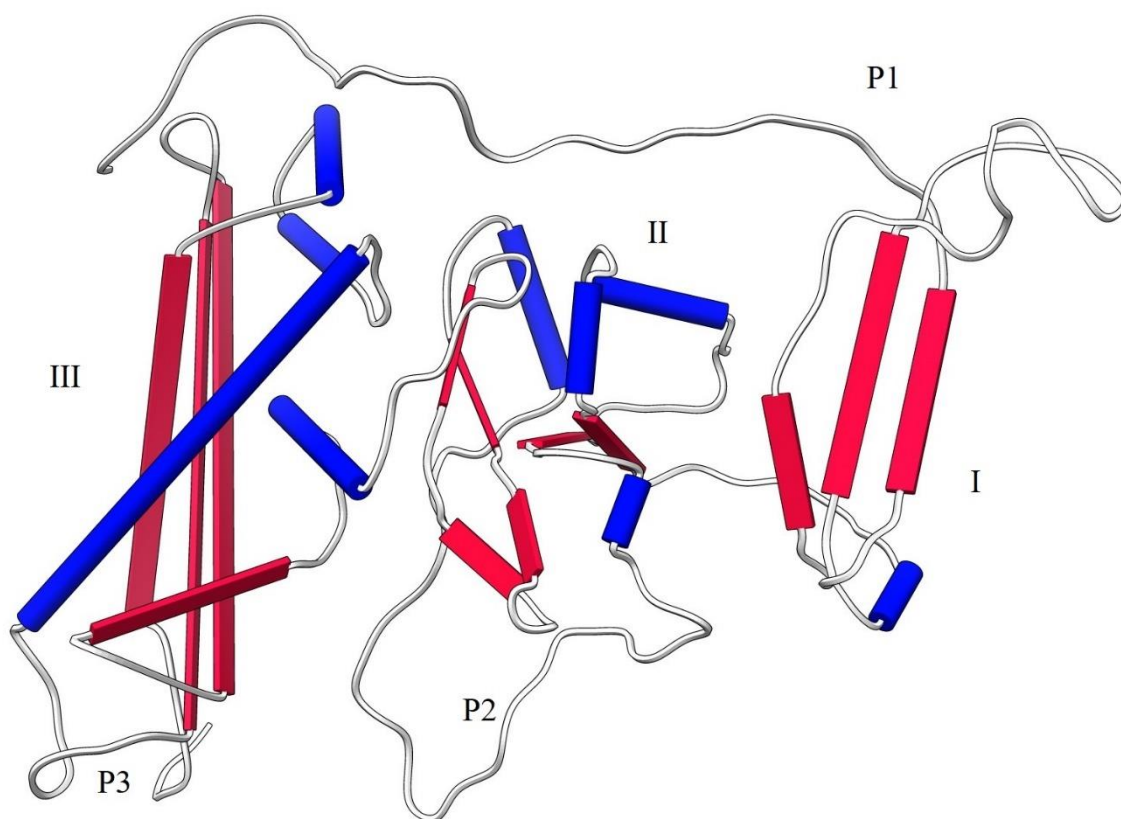
Badania nad strukturą krystaliczną homologa podjednostki C sugerują, że regiony II oraz III mogą być odpowiedzialne za interakcje z innymi podjednostkami receptora A lub

podjednostkami pozostałych receptorów kiełkowania. Elastyczny łącznik między domeną I oraz domeną II umożliwiać miałyby interakcje domen II i III z innymi podjednostkami, podczas gdy domena I byłaby zakotwiczona w wewnętrznej błonie przetrwalnika (Li, Setlow, Setlow i Hao, 2010).

Brak 24 reszt aminokwasowych w strukturze szablonu spowodował powstanie długiej nieuporządkowanej pętli obejmującej pierwsze reszty N-końca w modelu struktury trzeciorzędowej białka GerAC. Przewidywania struktury drugorzędowej wskazują, że region ten może przyjmować strukturę długiej α -helisy obejmującej reszty 3-15. Cysteina 18, która jest przypuszczalnie odpowiedzialna za łączenie podjednostki C z błoną, wydaje się być częścią nieuporządkowanego regionu. Natomiast helisa 3-15 jest prawdopodobnie fragmentem transbłonowym. Reszty 1-24 zostały usunięte z białka homologa podczas prób ustalenia jego struktury przestrzennej właśnie ze względu na błonowy charakter tego fragmentu.



Rysunek 13: Porównanie struktury drugorzędowej (2D) oraz struktury trzeciorzędowej (3D) białka GerAC. Kolorem czerwonym oznaczone struktury α -helikalne, natomiast kolorem zielonym β -kartki. Na zestawieniu widać wyraźny brak uporządkowania struktury na N-końcu białka w modelu struktury trzeciorzędowej oraz przewidywaną w tym miejscu helisę (reszty 3-15).



Rysunek 14: Struktura modelu białka GerAC pokazująca struktury drugorzędowe – helisy zaznaczone są kolorem niebieskim, β -kartki kolorem czerwonym. Domeny białka oraz pętle łączące te domeny są oznaczone tak, jak na rysunku 10.

6.1.3.3. Model białka GerKC – podjednostki C receptora GerK

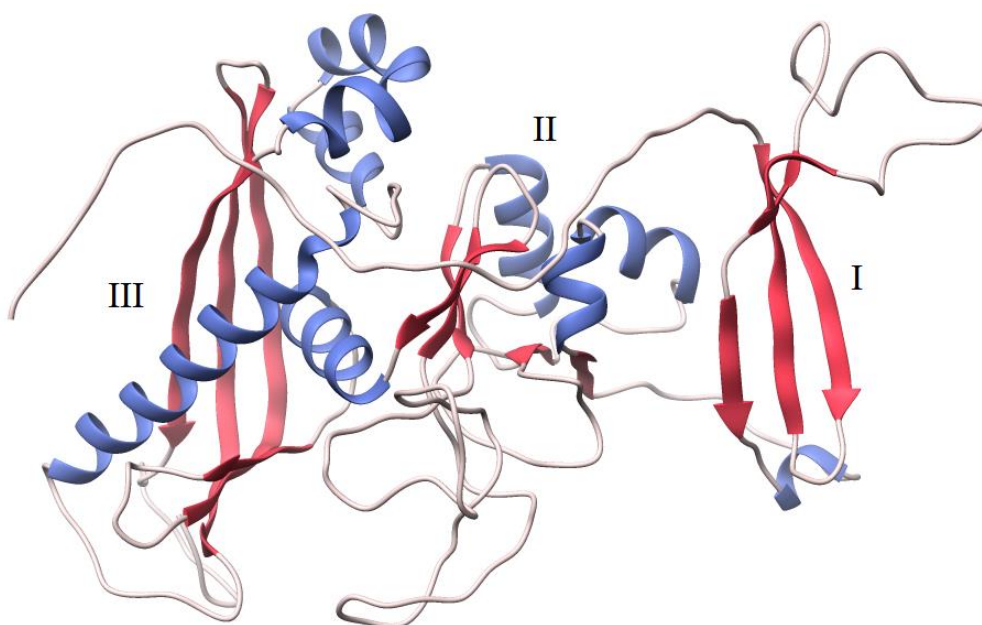
Receptor GerK jest receptorem, który inicjuje kiełkowanie przetrwalników *B. subtilis* w odpowiedzi na obecności mieszaniny AGFK w środowisku (L-asparaginan, D-glukoza, D-fruktoza oraz jony potasu).

Model białka GerKC, które jest podjednostką C receptora kiełkowania GerK, wykonałam na podstawie tego samego szablonu, który wykorzystałam do modelowania GerAC. Białko GerKC, podobnie jak omówione wcześniej białko C receptora GerA, są bliskimi homologami białka GerBC. Model podjednostki C receptora GerK został przeze mnie wykonany w ramach weryfikacji założenia, że homologiczne do siebie podjednostki receptorów kiełkowania będą przyjmować podobne struktury. Założenie to pojawiło się w

pracy badawczej opisującej strukturę GerBC (Li, Setlow, Setlow i Hao, 2010) i zostało poparte badaniami struktur drugorzędowych oraz zestawieniem sekwencji białek GerBC, GerAC oraz GerKC.

Model GerKC został przeze mnie wykonany według tego samego schematu, co model GerAC. Wyniki modelowania przedstawia rysunek 15. Model struktury trzeciorzędowej białka C receptora GerK charakteryzuje się znacznie większą ilością regionów nieuporządkowanych niż model białka GerAC. Zwłaszcza w obrębie domeny II łatwo zauważyć długą pętlę oraz mniej struktur drugorzędowych niż w modelu GerAC, czy samym szablonie.

Zestawienie sekwencji białka GerKC oraz białka GerBC (szablonu) pokazuje, że identyczność sekwencji plasuje się na poziomie 26%. Porównanie struktury szablonu z modelem białka docelowego za pomocą oceny RMSD dało wynik 0,59Å. Domeny I oraz III są zbudowane niemal identycznie jak w przypadku pozostałych dwóch białek C.



Rysunek 15: Model struktury białka GerKC. Na modelu zaznaczono domeny I, II oraz III. W obrębie domeny II widać duże nieuporządkowanie struktury. Domena ta składa się z trzech α -helis oraz pięciu β -katek. W domenie I pojawia się dodatkowa, krótka β -kartka umiejscowiona pomiędzy drugą a trzecią strukturą β tej domeny.

6.1.4. Podsumowanie — modele podjednostek receptora GerA

Dysponując niewielką ilością informacji o strukturze GerA podjęłam próbę uzyskania modeli strukturalnych podjednostek receptora. W trakcie prac w literaturze naukowej pojawiały się nowe informacje dotyczące białek i/lub homologów GerAA oraz GerAB, które na bieżąco brałam pod uwagę w pracach nad jakością modeli. Modele podjednostek wraz z danymi eksperymentalnymi dostępnymi w literaturze pozwoliły mi na zidentyfikowanie potencjalnych miejsc mutagenezy tych białek, jak również dostarczyły pewnych informacji strukturalnych na temat wpływu dokonanej mutacji na proces kiełkowania (Grela i inni, 2018).

Każda z trzech podjednostek GerA wykazuje wysoką homologię do odpowiednich podjednostek innych receptorów kiełkowania. Na tej podstawie można przypuszczać, że struktura będzie konserwowana pomiędzy homologami. Potwierdzeniem tego przypuszczenia może być podobieństwo struktur podjednostek C trzech różnych receptorów kiełkowania przetrwalników *B. subtilis* (GerA, GerB oraz GerK).

Uważa się, że zarówno GerAA, jak i GerAB mogą być odpowiedzialne, razem lub osobno, za interakcje z germinantami. Dysponując modelem strukturalnym tych białek mogłam przeprowadzić badania oparte na dokowaniu molekularnym L-alaniny i L-waliny w celu znalezienia potencjalnych miejsc wiązania. Wyniki dokowań opisałam w kolejnych punktach rozprawy.

6.2. Model białka GerD

Białko GerD jest lipoproteiną i nie wchodzi bezpośrednio w skład receptora kiełkowania GerA. Jego prawidłowe funkcjonowanie jest jednak związane z odpowiedzią receptora GerA na obecność germinantów w środowisku oraz inicjowanie procesu kiełkowania przetrwalników. Badania pokazują, że nieobecność białka GerD w sporze skutkuje dramatycznym spadkiem prędkości kiełkowania. Badacze spekulują, że białko GerD może być odpowiedzialne za gromadzenie się podjednostek lub całego receptora GerA w germinosomie. Białko D prawdopodobnie znajduje się po zewnętrznej stronie wewnętrznej błony przetrwalnika i, podobnie jak GerAC, jest połączone z błoną za pomocą kotwicy diacyloglicerolowej (Li i inni, 2014).

Model białka GerD przetrwalnika *B. subtilis* wykonałam na podstawie struktury jego bliskiego homologa — białka GerD *Geobacillus stearothermophilus*. Struktura homologa jest trimerem zbudowanym z trzech identycznych, równoległych łańcuchów polipeptydowych. Łańcuchy te są prawoskrętne i zbudowane są z 8, raczej krótkich, α -helis. Kryształ, na podstawie którego zdeterminowano strukturę GerD, składał się z dwóch trimerów, czyli sześciu molekuł (łańcuch A, reszty 69–178; łańcuch B, 63–179; łańcuch C, 64–176; łańcuch D, 60–179; łańcuch E, 61–175; łańcuch F, 66–180). Jak łatwo zauważyć, w modelu krystalicznym brakuje około 60 reszt aminokwasowych z N-końca białka. Dodatkowo, struktura tego białka została pozbawiona odcinka 181-195 na C-końcu (Li i inni, 2014).

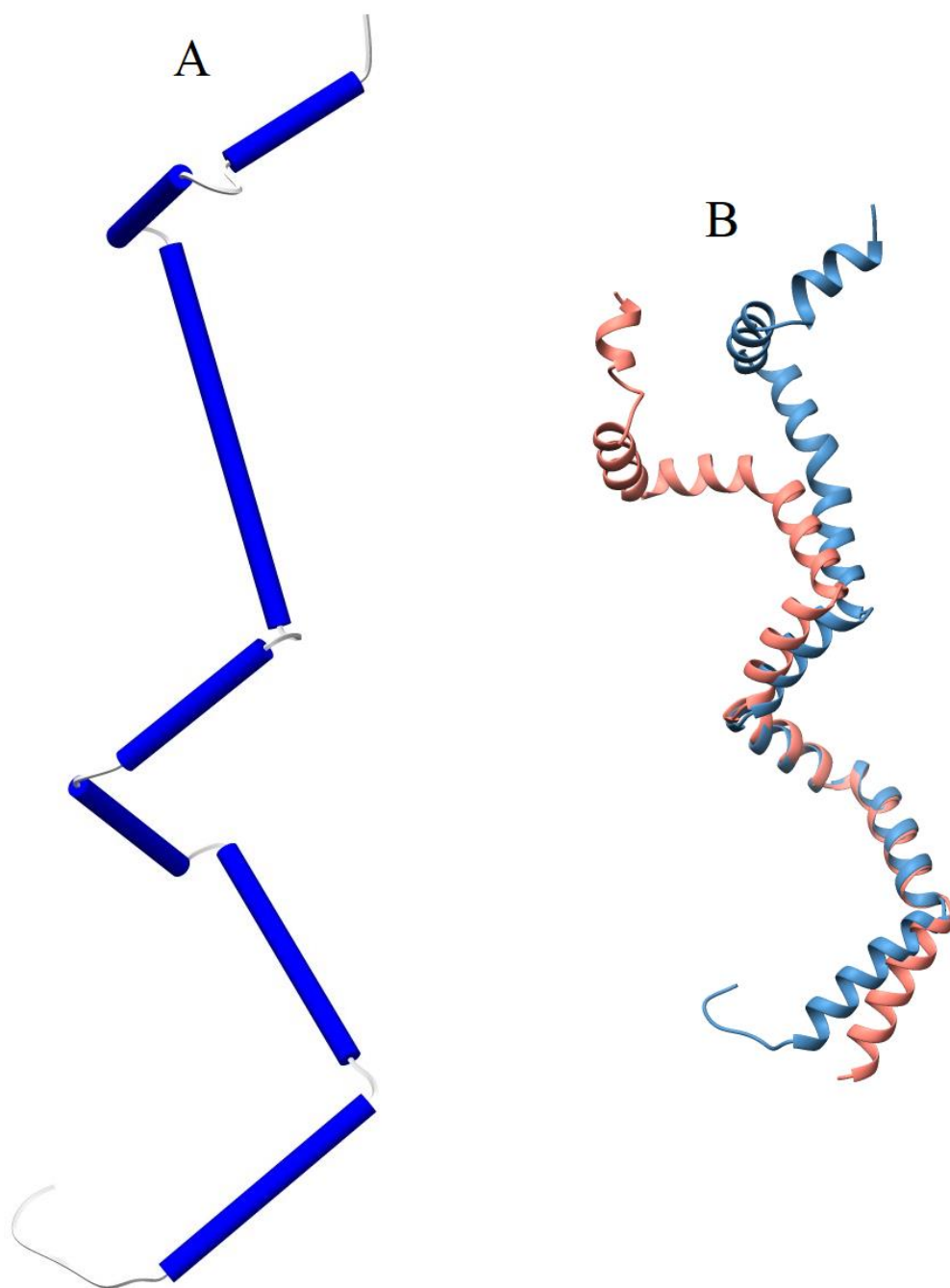
Podczas przygotowywania modelu struktury białka GerD *B. subtilis* wykorzystałam jeden z łańcuchów (łańcuch A) struktury krystalicznej do przeprowadzenia modelowania monomeru. Dodatkowo przewidziałam strukturę trimera — na podstawie dokowania symetrycznego modelu monomeru z wykorzystaniem oprogramowania SAM. Modelowanie, tak jak wspomniałam w rozdziale Materiały i Metody, zostało przeprowadzone za pomocą algorytmów oprogramowania Modeller.

6.2.1. Monomer GerD

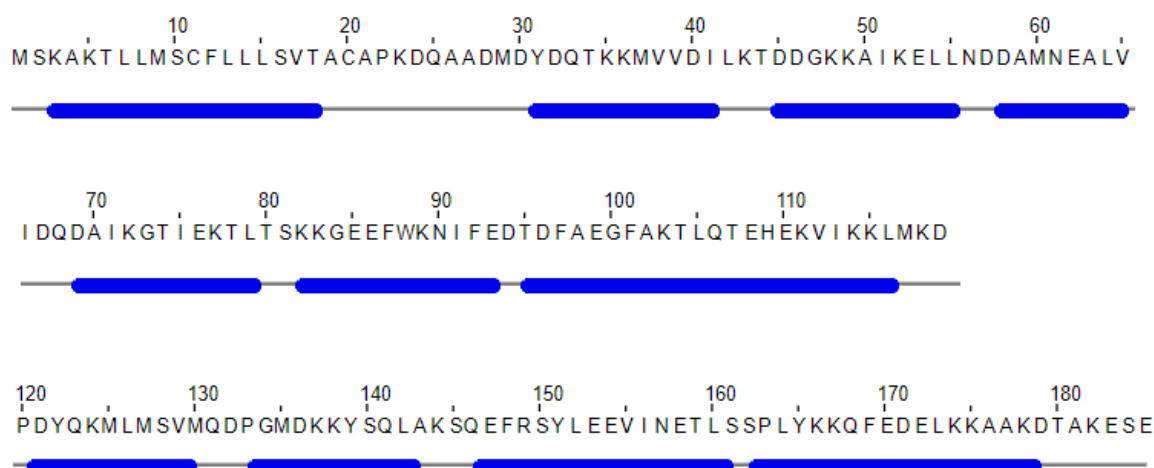
Białko GerD *B. subtilis* (białko docelowe) jest zbudowane ze 185 reszt aminokwasowych. Zestawienie sekwencji przy użyciu algorytmu ClustalW tego białka wraz z łańcuchem A białka homologicznego wykazało 49% identyczności tych sekwencji. Wartość ta wskazuje na wysoką homologię pomiędzy badanymi białkami, co czyni homologa idealnym szablonem do modelowania homologicznego. Modelowanie struktury trzeciorzędowej białka docelowego pozwoliło mi uzyskać model o wartości GA341 = 0,99636 oraz RMSD = 0,97Å. Obie wartości wskazują na wysokie podobieństwo struktury teoretycznej do struktury białka użytego jako szablon oraz wysoką jakość modelu.

Model struktury monomeru GerD składa się z 7 α -helis oraz długiego nieuporządkowanego regionu obejmującego reszty 1-70. Obecność tego regionu wynika z faktu, że w strukturze krystalicznej szablonu brakowało pierwszych 60 reszt aminokwasowych. Przewidywanie struktury drugorzędowej dla białka GerD pokazuje, że w miejscu nieuporządkowanego regionu struktura mogłaby przybrać formę 4 α -helis, które

obejmowałyby kolejno: reszty 2-18, 31-41, 45-55 oraz reszty 58-65 (rysunek 17). Te helisy prawdopodobnie odpowiadają za interakcje z błoną przetrwalnika (Li i inni, 2014).



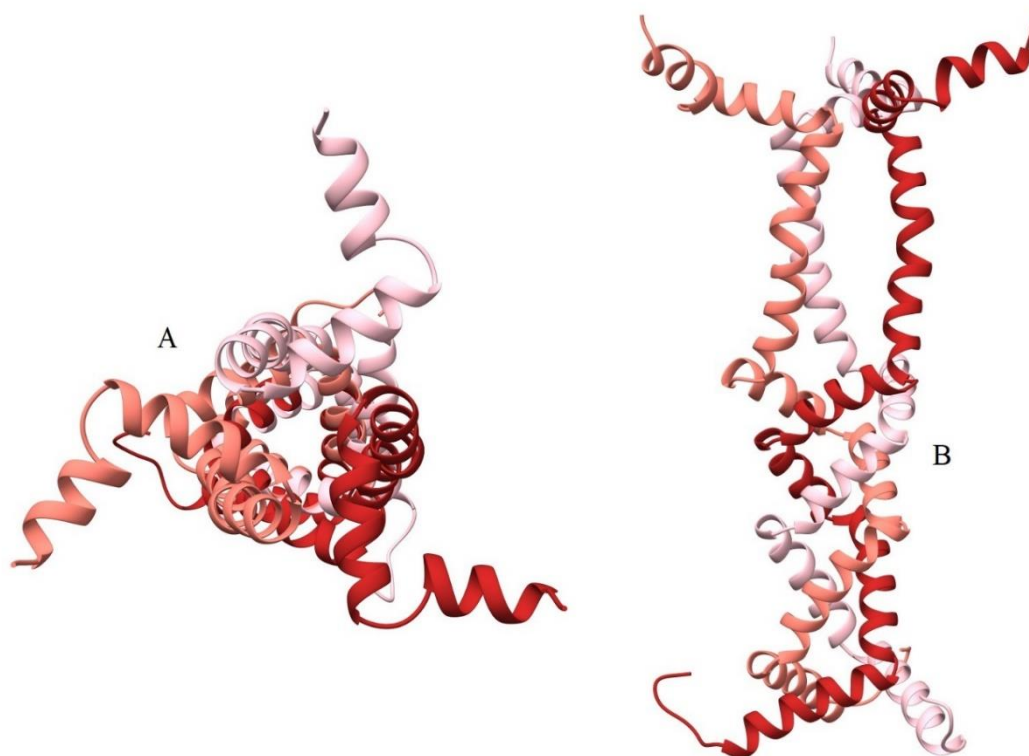
Rysunek 16: Model białka GerD bez pierwszych 60 reszt aminokwasowych. A) Model GerD pokazujący struktury drugorzędowe białka. GerD składa się z 7 helis (niebieski kolor). B) Porównanie modelu (zaznaczony kolorem niebieskim) oraz struktury szablónu. W strukturze szablónu (kolor pomarańczowy) brakuje ogonka na C-końcu, który jest dobrze widoczny w przypadku struktury modelu.



Rysunek 17: Wyniki przewidywania struktury drugorzędowej monomeru GerD. Przewidywania przeprowadzono za pomocą serwera JPred. Kolorem niebieskim zaznaczone są α -helisy. Wyraźnie widać cztery helisy przewidziane w rejonie 1-70.

6.2.2. Trimer GerD

Trimer GerD zbudowany jest z trzech równoległych monomerów. Jak już wspomniałam w punkcie 6.2, trimer białka GerD został zbudowany z wykorzystaniem metod dokowania bazujących na symetrii białek. Monomery zostały zadokowane do siebie z energią kinetyczną równą -431.95 kJ/mol. Model trimery białka D wykazuje się dużym podobieństwem do trimery szablonu (RMSD = $0,97\text{\AA}$). Do zbudowania modelu trimery wykorzystałam monomery, którym brakuje 70 reszt aminokwasowych z końca N. Być może ten brak znaczącej ilości reszt oraz interakcji pomiędzy ich atomami powoduje odgięcie helisy (H1) od środkowej osi trimery. To samo odchylenie pierwszej helisy obserwujemy w monomerze (rysunek 16).



Rysunek 18: Struktura trimeru GerD.

6.3. Mutacje w białku GerAA

Badania nad modelami struktur podjednostek receptora GerA oraz badania laboratoryjne nad funkcją białka GerAA dostarczyły wielu interesujących obserwacji. Na podstawie danych strukturalnych, informacji dostępnych w literaturze oraz obserwacji z badań prowadzonych w Zakładzie Bakteriologii Molekularnej Wydziału Biotechnologii UG oraz GUMed, wstępnie wytypowano trzy miejsca w sekwencji białka GerAA, które są niezwykle istotne dla poprawnego funkcjonowania receptora oraz poprawnej inicjacji procesu kiełkowania.

Wytypowane regiony i/lub reszty aminokwasowe zostały poddane mutagenzie, opracowane zostały modele strukturalne mutantów, a następnie badano różnice między modelami oraz zachowaniem i funkcjonowaniem przetrwalników zawierających różne warianty białka GerAA. Wszystkie modele mutantów zostały wykonane przeze mnie przy użyciu tych samych metod i parametrów, co podstawowe modele białka GerAA. Po

wprowadzeniu mutacji do sekwencji białka wykonywałam zestawienie sekwencji mutantów z sekwencjami szablonów (zestawienia były wykonywane tymi samymi metodami i algorytmami, które wykorzystano w przypadku podstawowych modeli), a następnie przeprowadzałam modelowanie przy użyciu oprogramowania Modeller.

W dalszej części rozdziału prezentuję wyniki tych eksperymentów ze szczególnym uwzględnieniem struktury białka GerAA oraz jego wariantów.

6.3.1. Pozycje 299 i 302

W swojej pracy badawczej doktor Anna Grela wykazała, że pozycje 299 oraz 302 białka GerAA są pozycjami istotnymi dla prawidłowego funkcjonowania receptora GerA (Grela i inni, 2018). Prowadzone przez nią badania uwzględniały dane strukturalne dostarczone przez modele podjednostki A receptora kiełkowania GerA oraz jej mutantów wykonane w ramach niniejszej pracy. W swoich badaniach dr Grela bazowała na *pierwszym* modelu GerAA, ponieważ w momencie prowadzenia eksperymentów nie były jeszcze dostępne dane strukturalne dotyczące niebłonowej części białek A, które pozwoliły na poprawę jakości przewidywań modelu. Niemniej, w dalszej części rozważań na temat wpływu reszt 299 oraz 302 na kiełkowanie przetrwalników oraz strukturę GerAA wzięłam pod uwagę również *drugi* model oraz różnice strukturalne między dwoma modelami.

Laboratoryjne szczepy *B. subtilis* wykazują pewne różnice w sekwencji białka GerAA w obrębie reszt 299 i 302 — są to różnice pojedynczych nukleotydów, które skutkują podstawieniem innych reszt aminokwasowych w tych miejscach.

Nasze badania skupiły się na sprawdzeniu różnic w kinetyce kiełkowania przetrwalników w dwóch szczepach używanych laboratoryjnie oraz mutantach zbudowanych na podstawie substytucji reszt aminokwasowych w pozycjach 299 oraz 302. Wszystkie warianty zostały przebadane dwutorowo: w laboratorium, pod kątem zdolności do kiełkowania przetrwalników i dynamiki odpowiedzi na germinanty, oraz *in silico*, pod kątem wpływu zmian na strukturę białka. Konstrukcja mutantów, badania *in vivo* oraz ich wyniki zostały szczegółowo opisane w pracy doktorskiej dr. Anny Greli (Grela A., 2019). W tym rozdziale skupimy się na różnicach strukturalnych mutantów oraz ich potencjalnym wpływie na kiełkowanie przetrwalników.

Przewidywanie struktury drugorzędowej białka GerAA pokazuje, że reszty 299 oraz 302 prawdopodobnie są częścią dużej α -helisy, która rozciąga się od reszty 274 do reszty

302 (rysunek 8). Dodatkowo, znacząca część tej helisy, jak pokazują przewidywania, stanowi region transbłonowy w domenie błonowej białka (w zależności od serwera: TMAP – reszty 280-302, TMPred – reszty 284-300 oraz TMHMM – reszty 279-301). Na podstawie tych danych można przyjąć założenie, że reszty 299 oraz 302 są strukturalnie istotne dla poprawnego funkcjonowania białka GerAA w receptorze. W ramach sprawdzenia wpływu tych dwóch reszt na strukturę całego białka wykonałam modele następujących mutantów: 299A302S; 299T302S; 299A302P oraz 299T302P.

Przewidywania struktury trzeciorzędowej opisane przez *pierwszy* model GerAA, pokazują, że warianty, w których w miejscu 302 podstawiono prolinę charakteryzują się krótszą helisą. Może to wynikać z faktu, że prolina jest tak zwanym „łamaczem helis”. Oznacza to, że prolina, ze względu na brak protonu amidowego, który tworzy wiązania wodorowe w obrębie helisy oraz obecność pierścienia, nie może być w pełni wbudowana w α -helisę. Zatem, obecność proliny w strukturze helikalnej często kończy się *złamaniem* lub skróceniem helisy. Przewidywania te zgadzałyby się z wynikami eksperymentów *in vivo*, które pokazują, że podstawienie P302 skutkuje brakiem kiełkowania przetrwalników.

Struktura trzeciorzędowa opisana przez *drugi* model podjednostki A pokazuje z kolei, że w każdym z wariantów helisa 274-302 kończy się przed resztą 300. W wariantach z alaniną w pozycji 299 helisa ta jest dłuższa o jedną resztę aminokwasową i sięga reszty 299. Helisa w tych modelach jest znacznie krótsza niż ta przewidziana przez algorytmy JPred do przewidywań drugorzędowych oraz *pierwszy* model GerAA. W takim przypadku, podczas budowania modelu zastosowałam więzy na reszty budujące helisę, które miałyby wskazać algorytmom Modellera, że w tym miejscu powinna znaleźć się helisa – ten zabieg nie wpłynął na wyniki i badana helisa rozciągała się od reszty 277 do 299.

Badania dr Anny Greli pokazują, że obecność proliny w miejscu 302 podjednostki A receptora skutkuje utratą funkcjonalności receptora — nie dochodzi do kiełkowania indukowanego przez L-alaninę i/lub L-walinę. W połączeniu z wynikami strukturalnymi oraz charakterystycznym zachowaniem proliny w strukturze helisy, możemy założyć, że struktura badanego regionu (273-302) jest istotna dla poprawnego funkcjonowania receptora. Może to wynikać z faktu, że najmniejsze zmiany w regionach hydrofobowych białek powodują zaburzenia w sposobie, w jakim białka wchodzi w interakcje z błoną komórkową. Reszta ta, jeżeli znajduje się w obrębie wewnętrznej błony przetrwalnika, tak jak sugerują to przewidywania regionów transbłonowych, może również pełnić ważną rolę w stabilizacji oddziaływań z innymi podjednostkami znajdującymi się w błonie.

Tabela 5: Długość helisy 273-302 w różnych wariantach oraz modelach GerAA. Tabela opisuje różnice strukturalne helisy w wynikach przewidywania struktury drugorzędowej oraz struktury trzeciorzędowej dwóch modeli GerAA.

Wariant	struktura 2D	pierwszy model	drugi model
299A/302S	274-302	273-302	277-299
299A/302P	273-302	274-300	277-299
299T/302S	274-302	274-302	277-298
299T/302P	273-302	274-300	277-298
299A/302A	273-302	x	277-299

6.3.2. Arginina 347

Ponieważ białko GerAA jest jedną z dwóch podjednostek typowanych jako białka wiążące germinanty, podczas swoich badań przeprowadziłam eksperymenty dokowania molekularnego L-alaniny oraz L-waliny do obu opisanych w tej pracy modeli. Dokładny opis wyników dokowań germinantów znajduje się w rozdziale 6.5.

Wyniki dokowań globalnych obu ligandów wskazują region 342-352 jako potencjalne miejsce wiązania germinantów. W regionie tym wiązała się duża liczba ligandów zarówno podczas dokowań do modelu białka GerAA. Arginina znajdująca się na pozycji 347 pojawiała się również jako miejsce interakcji podczas wstępnych badań nad dimerami podjednostki A w bardzo wczesnych etapach prac nad modelem tego białka. Dodatkowo, według danych uzyskanych podczas badania helis transbłonowych, region ten znajdować się może na powierzchni wewnętrznej błony przetrwalnika. Na podstawie tych danych postanowiliśmy zaprojektować eksperyment, który pozwoliłby na sprawdzenie hipotezy czy Arg347 jest resztą istotną dla prawidłowego funkcjonowania receptora GerA. Ze względu na silnie zasadowy charakter argininy zdecydowaliśmy na zmianę ładunku reszty 347 oraz sprawdzenie wpływu tej zmiany na funkcjonowanie receptora.

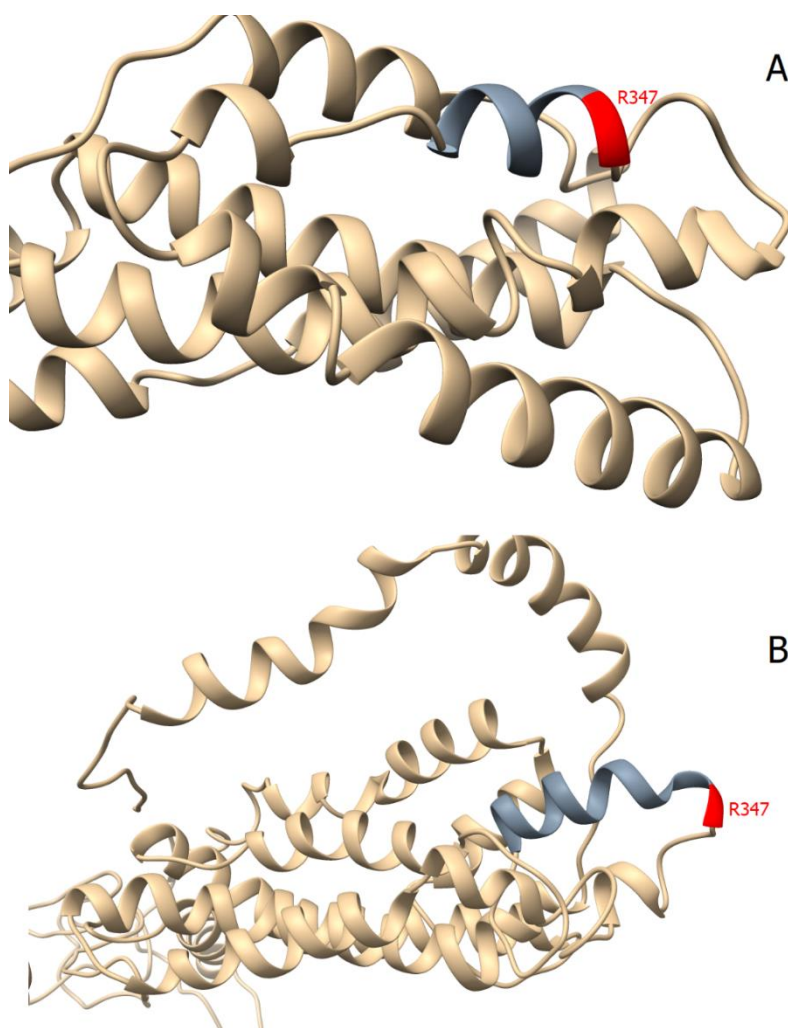
W swojej pracy doktorskiej doktor Anna Grela (Grela A. , 2019) opisuje eksperyment, którego celem było wstępne zbadanie istotności reszty Arg347 dla prawidłowego działania receptora (dane niepublikowane poza rozprawą dokorską doktor Anny Greli). Podczas eksperymentu wykorzystano cztery warianty białka GerAA, w których — zgodnie z założeniami zmiany ładunku — stopniowo zmniejszono zasadowość reszty

347. Pierwszy wariant GerAA w miejscu 347 zawierał argininę, kolejne warianty oparte były na substytucji argininy na kolejno: lizynę (347Lys), histydynę (347His) oraz alaninę (347Ala).

Analiza kinetyki kiełkowania receptora oparta na ilości uwalnianego DPA przeprowadzona przez dr Annę Greła, pokazuje, że wraz ze spadkiem zasadowości reszty 347 zmniejsza się tempo kiełkowania przetrwalników oraz ilość wydzielanego przez nie DPA. Spory z podstawioną lizyną w miejscu 347 kiełkowały wolniej niż typ dziki z argininą, natomiast w sporach zawierających histydynę lub alaninę w ogóle nie dochodziło do kiełkowania w obecności L-alaniny. Wyniki tej analizy wstępnie potwierdzają istotność pozycji 347 oraz jej zasadowości białka GerAA w procesie kiełkowania. Hipoteza ta wymaga jednak większej ilości badań, aby potwierdzić wysnute założenia oraz zbadać na ile jest to związane z wiązaniem germinantów, a na ile z innymi czynnikami istotnymi dla prawidłowego funkcjonowania receptora kiełkowania (Greła A. , 2019).

Według wykonanych przeze mnie przewidywań drugorzędowych arginina 347 znajduje się tuż za długą α -helisą obejmującą reszty 325-346. W *drugim* modelu GerAA reszta R347 znajduje się na końcu krótkiej helisy (helisa obejmuje miejsca 351-347). Natomiast w modelu *pierwszym* podjednostki A, tym który był dostępny podczas badań przeprowadzonych przez dr Annę Greła, arginina ta również znajduje się na końcu helisy (334-337). Według przewidywań regionów transbłonowych reszta ta jest fragmentem helisy błonowej, która w zależności od algorytmu obejmuje następujące reszty: TMAP 306-388, TMPred 373-392 oraz TMHMM 374-396.

Dalsze badania nad wpływem reszty 347 na strukturę białka GerAA lub jego oddziaływanie ze środowiskiem wymagają utworzenia modeli mutantów, nad którymi pracowała dr. Anna Greła oraz porównanie ich struktury, oddziaływań z germinantami oraz błoną.

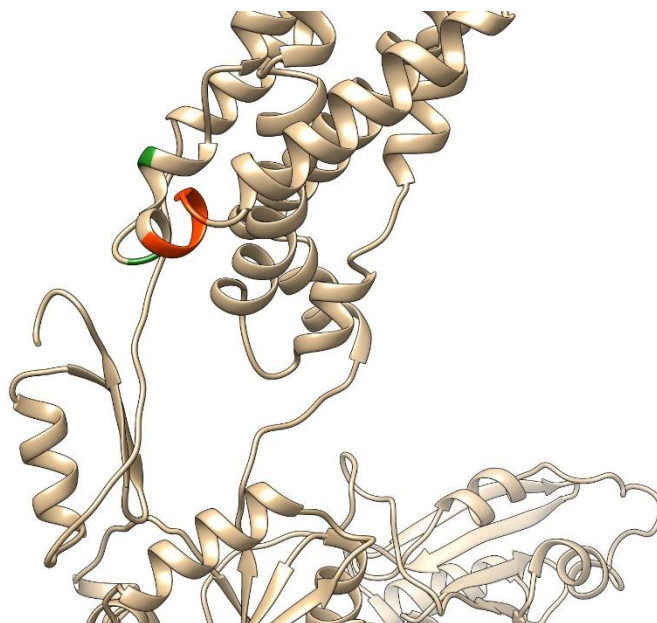


Rysunek 19: Strukturalne porównanie regionu występowania reszty 347 w dwóch modelach białka GerAA. A) Model *drugi* białka GerAA. B) *Pierwszy* model GerAA. Kolorem czerwonym zaznaczone są reszty 347 w obu modelach, natomiast kolorem szarym helisy, w których skład wchodzi R347.

6.3.3. Motyw PFPP

Motyw PFPP (motyw obejmujący reszty 324-327) znajdujący się w białku GerAA jest motywem zachowywanym wśród znanych podjednostek A receptorów kiełkowania. Region ten wraz z konserwatywnymi resztami E330 oraz E335 wydaje się być szczególnie ważny dla funkcjonowania receptora oraz dla struktury białka GerAA. Badania pokazują, że wprowadzenie mutacji w tym regionie skutkuje różnymi fenotypami receptora, m.in.: zwiększoną wrażliwością na germinanty i brakiem klasteryzacji (nieobecna podjednostka C przy mutacjach w miejscu E330 i E335). W swojej pracy Mongkolthanasakul i inni (Mongkolthanasakul, Cooper, Mawer, Allan i Moir, 2011), badając wpływ różnych mutacji (w tym tych w motywie PFPP oraz w pozycjach E330 i E335) na działanie receptora GerA,

sugerują, że region PFPP reprezentować może nietypową helisę błonową lub pętlę łączącą helisy błonowe, która byłaby wyeksponowana poza wewnętrzną błonę przetrwalnika.



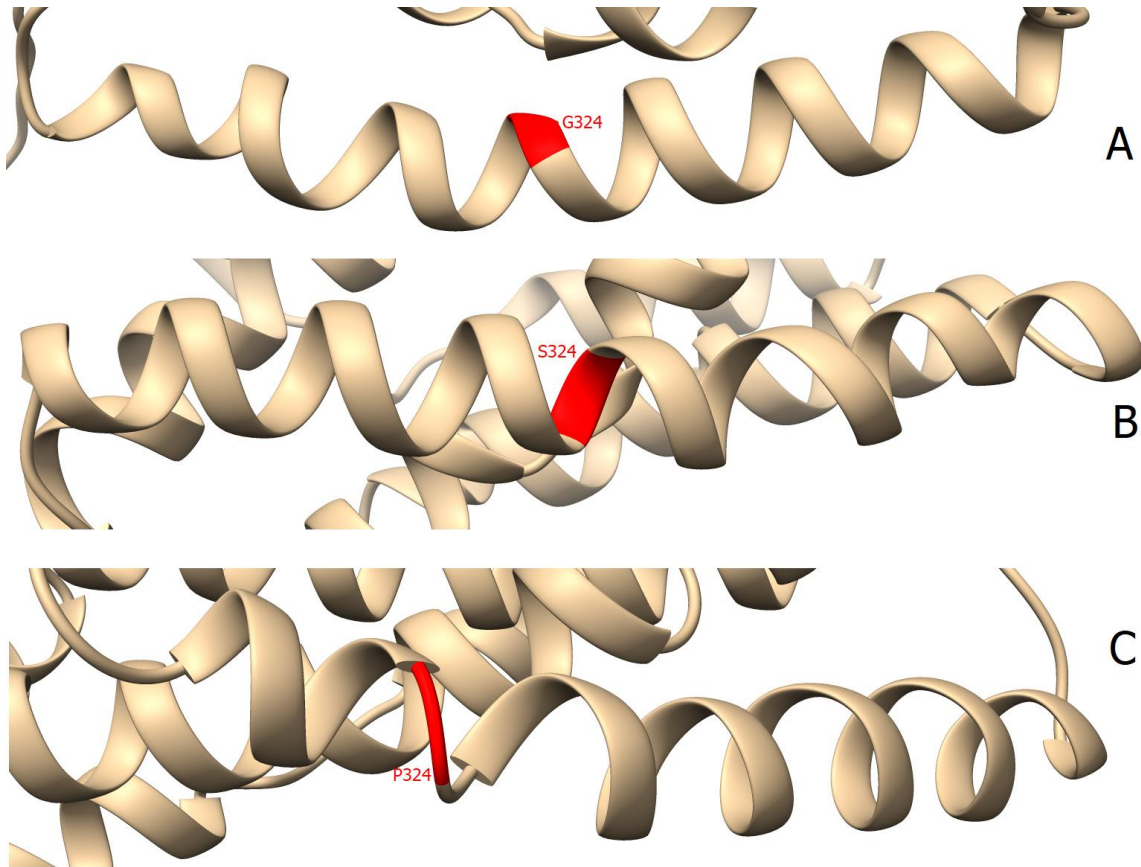
Rysunek 20: Fragment modelu struktury GerAA z zaznaczonym regionem PFPP (kolor czerwony) oraz resztami 330 i 335 (kolor zielony).

Utworzony przeze mnie model białka GerAA (model *drugi*) pokazuje, że początek motywu PFPP — reszta P324 oraz poprzedzająca ją V323 — buduje pętlę łączącą dwie helisy. Reszty 325-327 z kolei wchodzi w skład helisy. Cały motyw jest natomiast rejonem łączącym dwie helisy błonowe — jak pokazują wyniki przewidywań regionów transbłonowych (tabela 3) — i prawdopodobnie znajduje się na powierzchni błony po jej wewnętrznej stronie.

W związku z wynikami wyżej wspomnianych prac badawczych wykonałam modelowanie struktury dwóch mutantów, w których dokonałam substytucji proliny w miejscu 324 na alaninę oraz na glicynę w celu zweryfikowania wpływu mutacji w tym regionie na strukturę białka GerAA. Wyniki modelowania mutantów osiągnęły znacznie słabsze oceny niż modele GerAA bez mutacji, co jest niezwykle ciekawe, biorąc pod uwagę, że była to zmiana punktowa, skupiona na jednej reszcie aminokwasowej.

W obu modelach mutantów można zaobserwować, że motyw PFPP wchodzi w skład dużej α -helisy. Pętla, która obejmowała reszty 323 oraz 324 i łączyła dwie mniejsze helisy nie występuje w modelach mutantów. Prawdopodobnie obecność proliny — znanego

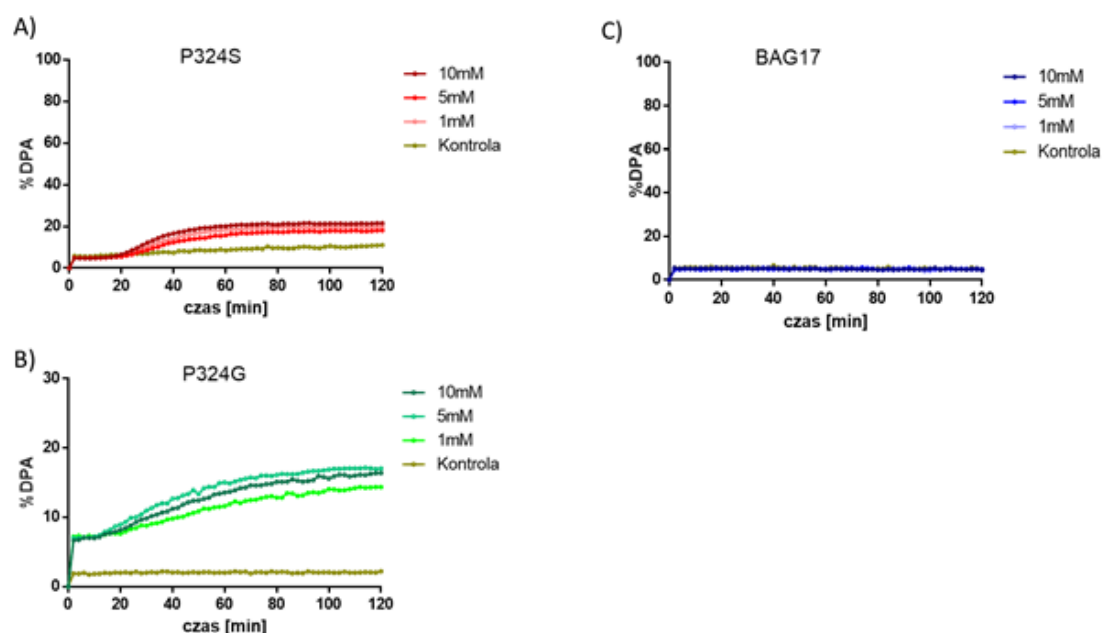
łamacza helis (Pace i Scholtz, 1998) — w sekwencji białka dzikiego typu zapewniała strukturalne oddzielenie dwóch helis.



Rysunek 21: Porównanie struktur wariantów GerAA A) Mutant P324G — w tym wariantcie reszta 324 jest częścią dłuższej α -helisy (310-336). B) Mutant P324S, podobnie jak w punkcie A, reszta 324 jest wbudowana w helisę 310-337. C) Struktura wariantu dzikiego GerAA z proliną w miejscu 324. Prolina wraz z waliną 323 budują łącznik między dwiema α -helisami.

Badania przeprowadzone przez dr Martę Hubisz w Zakładzie Bakteriologii Molekularnej Wydziału Biotechnologii Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego nad motywem PFPP oraz jego mutacjami sugerują, że receptory z mutacjami P324S i P324G w podjednostce A mogą kiełkować w obecności D-alaniny (dane niepublikowane), która jest opisywana w literaturze jako silny inhibitor kiełkowania. P324G wykazuje dużo wyższe wydzielanie DPA w obecności D-alaniny niż P234S oraz wariant niezmieniony (rysunek 22). Dodatkowo receptory z mutacjami kiełkowały szybciej w niskich stężeniach L-alaniny niż wariant szczepu dzikiego.

Opisane mutacje w motywie PFPP zaburzą strukturę regionu, który stanowi łącznik między fragmentami transbłonowymi i znajduje się blisko krawędzi błony. Opisany fragment jest wyeksponowany poza wewnętrzną błonę przetrwalnika. Informacje o zmianach strukturalnych mutantów w zestawieniu z wynikami badań dr Marty Hubisz mogą sugerować, że zmiany w obrębie konserwowanego motywu PFPP mogą powodować zaburzenia w kieszeni wiążącej germinanty, co pozwala na kiełkowanie receptora w obecności D-alaniny. Dodatkowo zmiany struktury mogą powodować zaburzenia w interakcji białka z wewnętrzną błoną przetrwalnika, co również może wpływać na interakcje z germinantami oraz inhibitorem kiełkowania. Na tej podstawie można wysnuć wniosek, że przedstawione w tej pracy oraz w pracy dr Marty Hubisz przetrwalniki z substytucjami reszty 324 są bardziej wrażliwe na germinanty w procesie inicjowania kiełkowania spor.



Rysunek 22: Wykres przedstawiający uwalnianie DPA kiełkujących szczepów A)P324S B)P324G oraz C) Szczepu dzikiego z proliną w miejscu 324 (BAG17). Wykres został wykonany przez dr Martę Hubisz i jest częścią jej pracy doktorskiej.

6.4. Modele receptora

Zbudowanie modeli struktur podjednostek receptora GerA jest pierwszym krokiem na drodze do opracowania struktury całego receptora. Wykonane przeze mnie modelowanie homologiczne podjednostek receptora oraz wszelkie dostępne w literaturze dane dostarczyły wielu informacji strukturalnych niezbędnych do dalszych badań nad całym receptorem. Pomimo tego, pozostaje jeszcze kilka, bardzo istotnych, otwartych pytań, na które nie znamy odpowiedzi. Jednym z najważniejszych jest pytanie o stechiometrię podjednostek w receptorze. Z dostępnych danych wynika, że stosunek podjednostek w receptorze wynosi 1:1:1. Oznacza to tyle, że najprawdopodobniej receptor zbudowany jest z jednakowej liczby podjednostek i wciąż niewiele mówi o tym, ile tych podjednostek w receptorze rzeczywiście jest.

Na podstawie podobieństwa do wzorców użytych w modelowaniu struktur podjednostek B i C spekulujemy, że, w wysokim stężeniu, podjednostki receptora mogą tworzyć dimery, które następnie budowałyby funkcjonalny receptor. Białka AdiC należące do APC, użyte jako szablony w modelowaniu GerAB, tworzą dimery z miejscem wiążącym ligand na styku obu monomerów (Gao i inni, 2010). GerBC, bliski homolog podjednostki C, w postaci krystalicznej buduje dimer, natomiast w roztworze zachowuje strukturę monomeru, co pozwala przypuszczać, że wysokie stężenie białka może prowadzić do tworzenia dimerów (Li, Setlow, Setlow i Hao, 2010).

Biorąc pod uwagę skłonności do dimeryzacji szablonów użytych podczas modelowania, mogę przypuszczać, że podjednostki GerA będą tworzyć dimery. Dodatkowym argumentem przemawiającym za tym założeniem jest podobieństwo procesu klastrowania (grupowania) receptorów kiełkowania do grupowania bakteryjnych matryc chemotaksji (Sherman, 1982; Sourjik i Armitage, 2010). Zatem w dalszych częściach prac nad modelem receptora GerA przyjąłm jako jedno z założeń, że receptor może tworzyć trimer dimerów.

Pomimo wyżej wymienionych przesłanek sugerujących trimer dimerów, pod uwagę wzięłam także prawdopodobieństwo tego, że receptor mogą budować monomery podjednostek. Taki model receptora został, obok receptora zbudowanego z dimerów, wzięty pod uwagę w badaniach nad miejscami wiązania germinantów oraz orientacji receptora w błonie.

6.4.1. Wyniki przewidywania miejsc interakcji białko-białko

Przewidywanie miejsc interakcji białko-białko dla każdej podjednostki przeprowadziłam za pomocą serwera meta-PPISP oraz serwera cons-PPISP. Wyniki przedstawione zostały w tabeli poniżej.

Tabela 6: Wyniki przewidywań miejsc interakcji poszczególnych podjednostek receptora GerA z innymi białkami receptora lub (potencjalnie) innymi białkami biorącymi udział w procesie kiełkowania przetrwalnika. Kolorem zielonym zaznaczono reszty, które oddziałują między monomerami w kompleksie dimeru i/lub w kompleksie receptora.

Białko	Serwer	Reszty aminokwasowe
GerAA	meta-PPISP	V140, I143, D144, I279, R280, R283, F284, S286, I287, I289, T290, L291, F292, L293, S294, I296, Y297, I298, A299, L300, V301, L404, S411, L417, Y418, G419, I420, L422, F423, L425, V426, V427, Y428, T249, H430, L450
GerAA	cons-PPISP	F238, R241, L245, V246, D247, S248, S249, L320-L348, I356-V359, V363, Y395, L399, F409, S408
GerAB	meta-PPISP	I127, C133, G134, L161, I170, I237, W280, I281, F285, V289, Y320, F321, W324, P325, D326, V331, M333, Y334, S335, D336, Y337, L338, G339, I341, F342, V343, S344, L345, L352
GerAB	cons-PPISP	N40, L137-L148, I170, F172, I173, N174, L176-L180, L184
GerAC	meta-PPISP	R4, I5, L6, C7, M8, F9, I10, C11, T12, L13, L14, L15, S16, G17, C18, W19, D20, S21, D261, G262, D348, Y349, F350, S351, N352
GerAC	cons-PPISP	I5-W19, H94, L95, I98, I115-T124

6.4.2. Dimery podjednostek

Dimeryzacja podjednostek receptora GerA była pierwszym podjętym przeze mnie założeniem w kwestii budowy receptora. Wszystkie dimery zbudowałam z wykorzystaniem metod dokowania typu białko-białko. W trakcie przygotowywania modeli dimerów wykorzystywałam następujące oprogramowania: Sam, GrammX oraz ClusPro, a do dalszych badań wybrane zostały najlepsze kompleksy. Narzędzia służące do dokowania zostały dokładniej opisane w rozdziale 5.7.3 *Dokowania białko-białko*.

6.4.2.1. *Dimer GerAA*

Model dimeru GerAA powstał przy użyciu dokowania symetrycznego na serwerze ClusPro i, jak pokazuje rysunek 23, oddziaływania między monomerami skupiają się w błonowej części białka. Kompleks został wybrany spośród wielu innych kompleksów utworzonych w serwerze ClusPro oraz za pomocą dokowania symetrycznego w programie Sam. Kryteriami wyboru, które brałam pod uwagę były: energia potrzebna do zadokowania monomerów do siebie oraz ułożenie fragmentów transbłonowych.

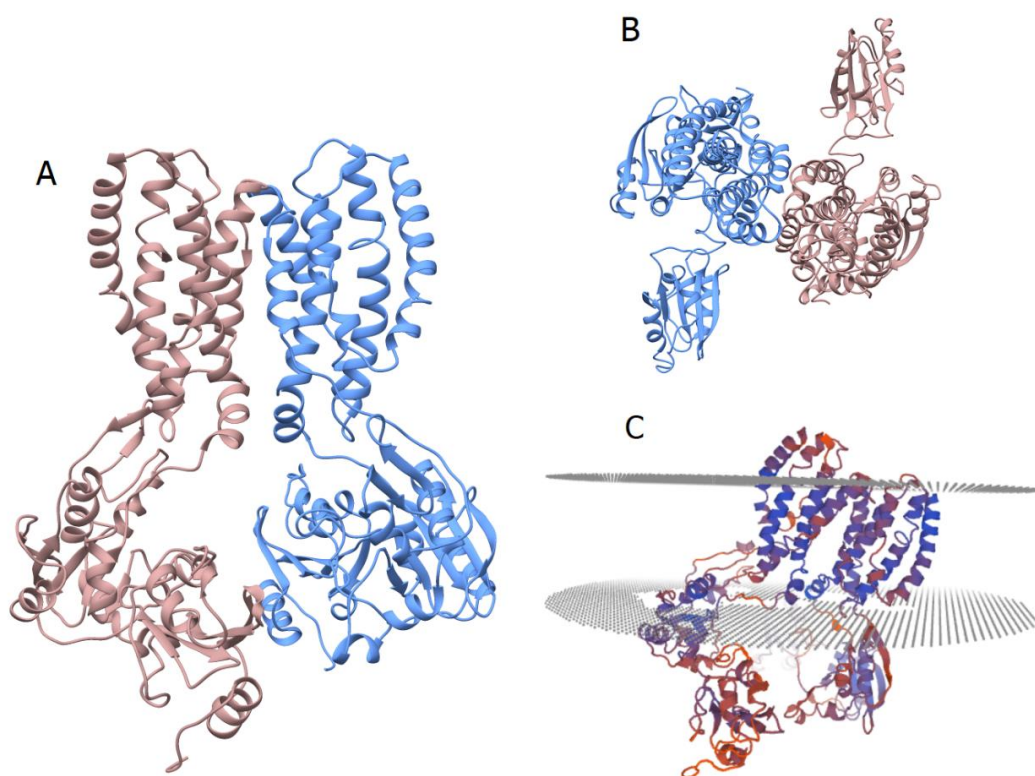
Monomery są ze sobą połączone za pomocą wiązań wodorowych oraz oddziaływań hydrofobowych, a sama interakcja między nimi skupia się głównie w domenie błonowej monomerów. 17 reszt aminokwasowych podjednostki A oraz 17 aa z podjednostki B jest zaangażowanych w oddziaływania hydrofobowe pomiędzy monomerami. Dodatkowo pomiędzy monomerami tworzy się 7 wiązań wodorowych oraz jedno wiązanie kowalencyjne. Wiązanie kowalencyjne, które pojawiło się podczas analizy wyników jest najprawdopodobniej artefaktem oprogramowania LigPlot. W metodach dokowania wiązania tego typu nie powinny się tworzyć — do utworzenia wiązania kowalencyjnego niezbędne byłoby zajście jakiejś reakcji enzymatycznej. Rozkład reszt aminokwasowych biorących udział w interakcjach pomiędzy monomerami przedstawiłam na rysunku 24.

Pro349, Arg402 oraz Tyr418 tworzą wiązania wodorowe w obu monomerach. Wiązanie Pro349-Tyr418 jest symetryczne, czyli prolina 349 z monomeru A wiąże się z tyrozyną monomeru B i odwrotnie (Pro349B z Tyr418A). Natomiast arginina z pozycji 402 monomeru A tworzy wiązanie wodorowe z Gln434, a arginina 402 B oddziałuje z Tyr428. Dodatkowo reszta Glu271 z monomeru B tworzy wiązania z resztami Ser277 i Arg280 monomeru A, przy czym wiązanie z Arg280 jest wiązaniem kowalencyjnym. Długość wiązań wodorowych zaobserwowanych w oddziaływaniach między monomerami nie przekracza 2,88Å.

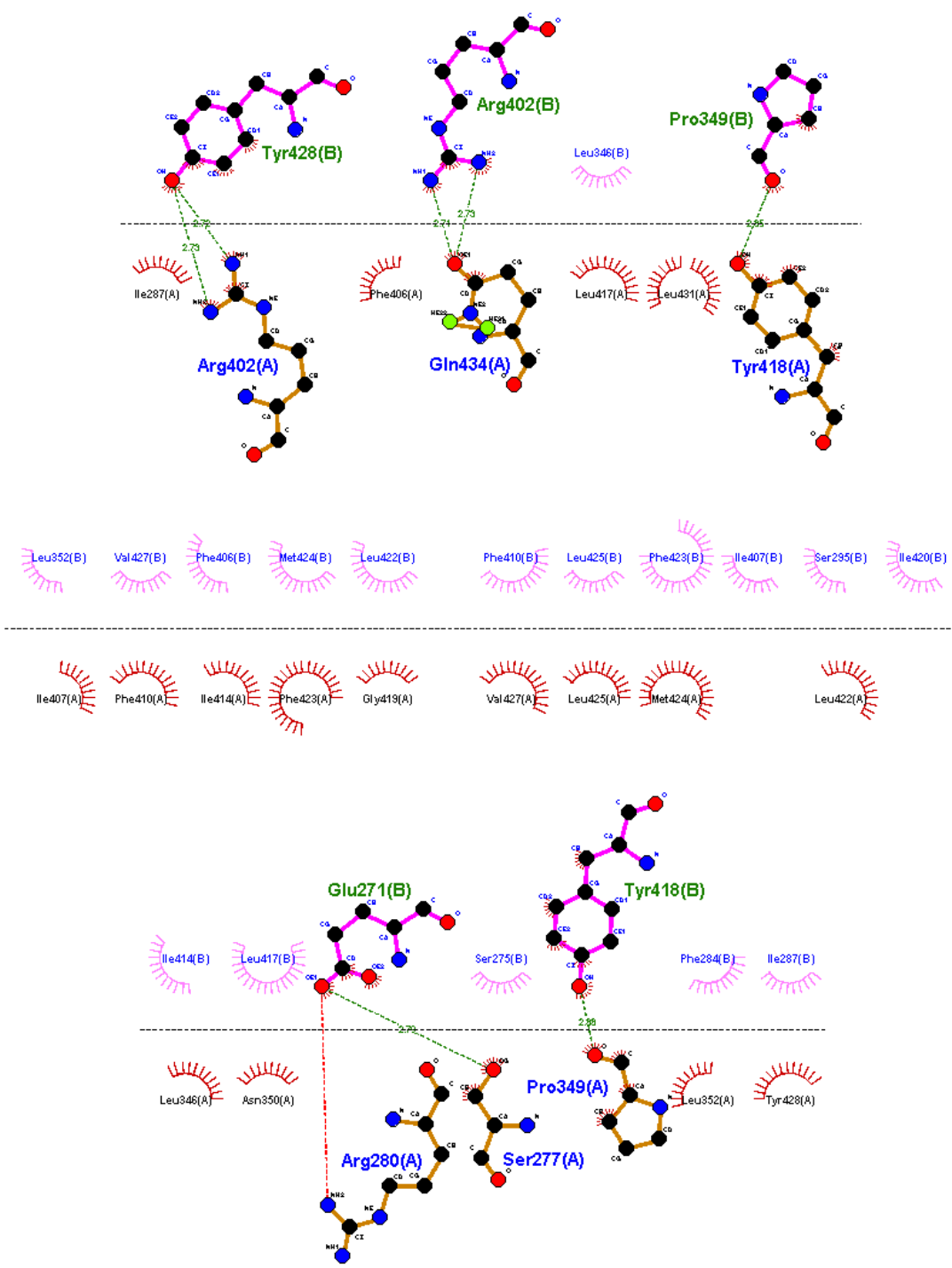
Badania pokazują, że substytucja proliny 349, zaangażowanej w oddziaływania między monomerami, na serynę skutkuje fenotypem przetrwalnika, w którym obserwuje się wolniejsze kiełkowanie. Co więcej, arginina 402 może być pierwszą resztą aminokwasową budującą helisę błonową lub być fragmentem pętli pomiędzy helisami błonowymi, która jest umieszczona po zewnętrznej stronie błony (tabela 3). Te same badania, które opisują wpływ mutacji P349S na działanie receptora, sugerują, że ta pętla pomiędzy helisami błonowymi może być niezwykle istotna dla funkcjonowania receptora, a zmiany strukturalne w jej obrębie skutkować utratą funkcjonalności receptora (w cytowanej pracy badawczej

przebadano pod tym kątem mutację S400F (Mongkoltharuk, Cooper, Mawer, Allan i Moir, 2011)).

Na podstawie wyników dokowania monomerów GerAA oraz informacji o zachowaniu mutantów w procesie kiełkowania przetrwalnika, możemy wysnuć wniosek, że opisywane mutacje mogą zakłócać tworzenie się dimeru poprzez całkowite uniemożliwienie dimeryzacji lub poprzez osłabianie oddziaływania monomerów, a tym samym destabilizację kompleksu.



Rysunek 23: Model dimery GerAA. Łańcuch A jest zaznaczony kolorem niebieskim, natomiast łańcuch B – kolorem różowym. A) Kompleks dimeru GerAA B) Widok na kompleks z góry. C) Orientacja dimeru w błonie obliczona narzędziem QMbrane – regiony oznaczone kolorem niebieskim to regiony o wysokiej jakości przewidywania, natomiast kolor czerwony pokazuje regiony niskiej jakości. Błona oznaczona jest kolorem szarym.



Rysunek 24: Rozkład reszt biorących udział w oddziaływaniach monomerów GerAA. Oddziaływania wygenerowane zostały za pomocą programu LigPlot+. Kolorem zielonym oznaczone są wiązania wodorowe, natomiast kolorem czerwonym – wiązanie kowalencyjne między atomami konkretnych reszt aminokwasowych, które prawdopodobnie jest artefaktem oprogramowania.

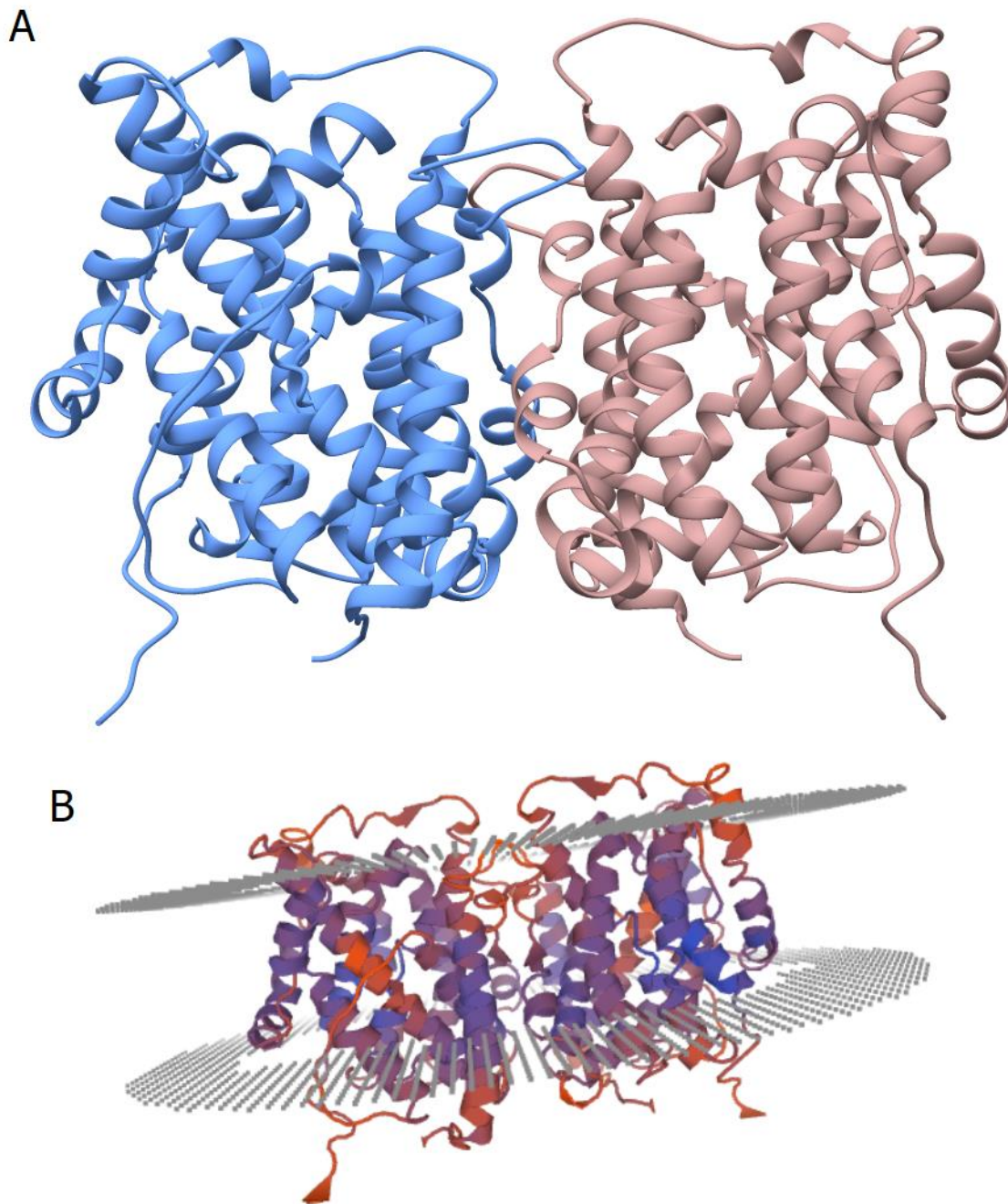
6.4.2.2. *Dimer GerAB*

Podjednostka B receptora kiełkowania GerA, jak już wielokrotnie wspomniano, wykazuje niewielką homologię do rodziny transporterów APC. Białka z tej rodziny formują dimery oraz są odpowiedzialne za wiązanie ligandów, a miejsce wiązania cząsteczek liganda znajduje się na styku dwóch monomerów. Na tej podstawie można zaryzykować twierdzenie, że GerAB również może formować dimery oraz, co wielokrotnie podkreśla się w literaturze, odpowiadać za interakcje z germinantami.

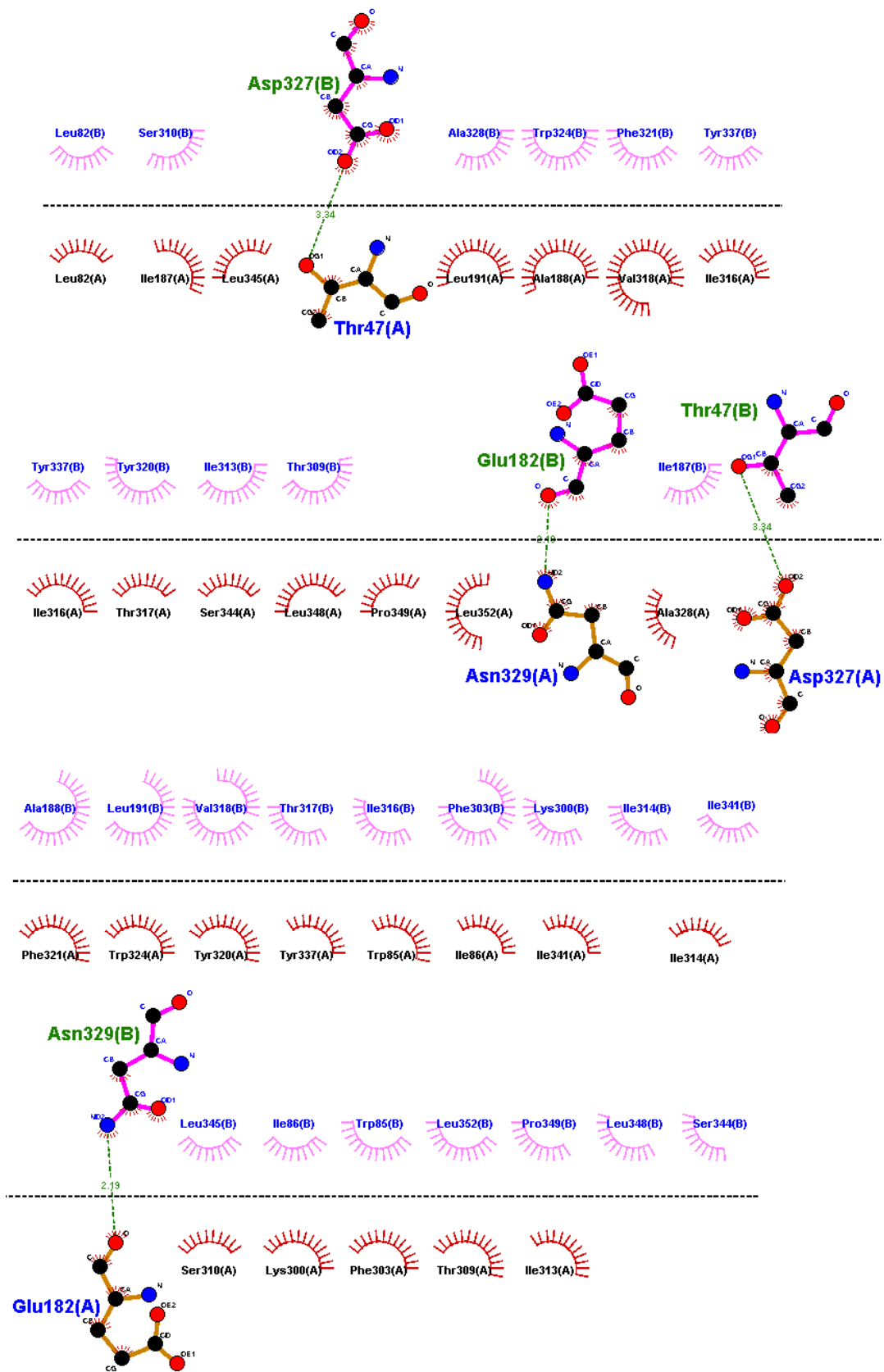
Dimer GerAB zbudowałam przy użyciu metod dokowania symetrycznego (symetria C2) zaimplementowanych w programie Sam 2.0.4. Energia wiązania kompleksu wybranego do dalszych badań wynosi -898,56 kJ/mol. Model dimeru oraz jego potencjalna orientacja w błonie pokazane są na rysunku 25. Szczegółowe obliczenia związane z orientacją GerAB w błonie zostały opisane w rozdziale *Orientacja białek receptora GerA w błonie*.

Monomery GerAB oddziałują ze sobą głównie za pomocą oddziaływań hydrofobowych — 26 reszt aminokwasowych monomeru A oraz 26 reszt monomeru B bierze udział w tych oddziaływaniach. Ponadto dimer jest stabilizowany czterema wiązaniami wodorowymi. Wiązania te utworzyły się pomiędzy następującymi resztami aminokwasowymi: Asp327(B)-Thr47(A), Glu182(B)-Asn329(A), Thr47(B)-Asp327(A) oraz Asn239(B)-Glu182(A).

Badania pokazują, że podstawienie alaniny w miejscu kwasu asparaginowego 327 daje mutanty, które mogą inicjować kiełkowanie w niższym stężeniu alaniny niż spory typu dzikiego (Cooper i Moir, 2011). Być może zmiana reszty aminokwasowej o odczynie kwaśnym na resztę aminokwasową o odczynie obojętnym, do tego z mniejszym łańcuchem bocznym, powoduje zwiększenie stabilności dimerów.



Rysunek 25: Model dimeru GerAB. A) Monomer A zaznaczony jest kolorem niebieskim, natomiast monomer B kolorem różowym. B) Orientacja modelu dimeru w błonie obliczona za pomocą narzędzia QMbrane. Kolor niebieski wskazuje regiony o dobrej jakości przewidywań, natomiast kolor czerwony – regiony gorszej jakości. Kolorem szarym oznaczone są granice błony.

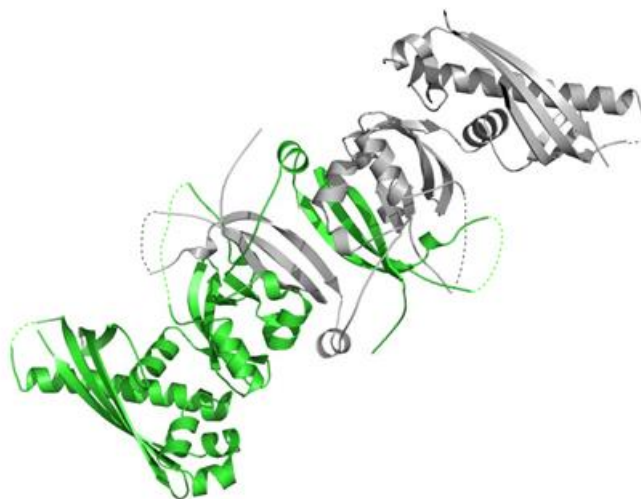


Rysunek 26: Rozkład wiązań wodorowych oraz oddziaływań hydrofobowych w dimerze GerAB.

6.4.2.3. *Dimer GerAC*

Homolog białka C receptora GerA, białko GerBC, w kryształach tworzy dimer (Li, Setlow, Setlow i Hao, 2010), natomiast w roztworze pozostaje monomerycznym. Może to świadczyć o tym, że przy dużym stężeniu podjednostki C receptorów kiełkowania wykazują skłonności do dimeryzacji. Zatem, w momencie inicjacji procesu kiełkowania, gdy w germinosomie liczba białek receptorów podjednostki C jest bardzo duża (a co za tym idzie, według stechiometrii 1:1:1, inne podjednostki), GerBC może tworzyć dimery.

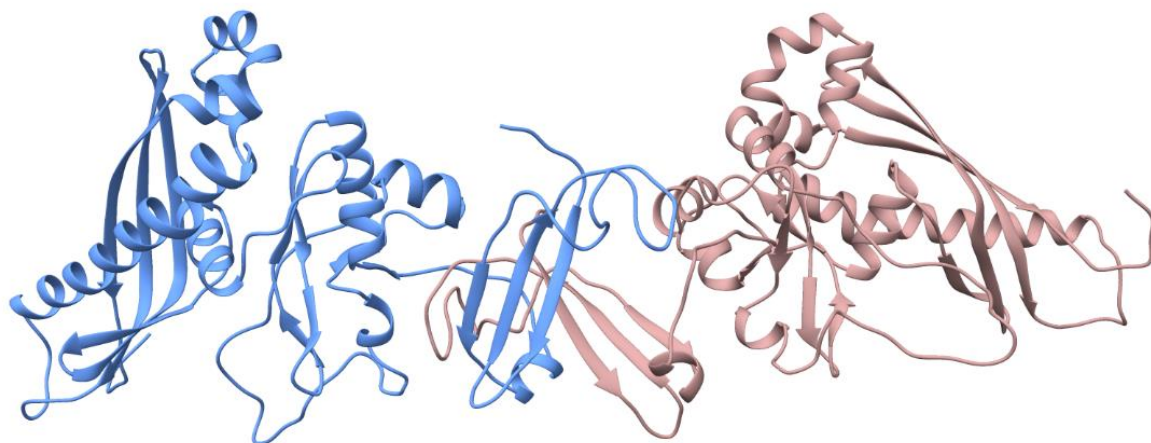
W dimerze GerBC za oddziaływanie monomerów odpowiadają domeny I oraz II. Domena I monomeru A znajduje się pomiędzy domenami I a II monomeru B (i odwrotnie), na co prawdopodobnie pozwala elastyczna pętla łącząca domenę I oraz II.



Rysunek 27: Struktura dimeru GerBC. Kolorem zielonym oznaczony jest monomer A, a kolorem szarym – monomer B. Struktura oraz ilustracja pochodzą z publikacji (Li, Setlow, Setlow i Hao, 2010).

Dimer GerAC zbudowałam z wykorzystaniem oprogramowania GrammX. Podczas dokowania monomerów białka C dużą przeszkodą okazała się pętla obejmująca pierwsze 24 reszty aminokwasowe N-końca monomerów, dlatego została ona usunięta na potrzeby zbudowania dimeru jak najbardziej podobnego do dimeru homologa (GerBC). Usunięcie pętli z modelu sprawiło, że modele monomerów zadokowały się do siebie poprzez przeplatanie się domen I. Być może efekt dimeryzacji obserwowany w kryształach wynika właśnie z tego, że w strukturze GerBC brakuje 24 reszt aminokwasowych z N-końca białka

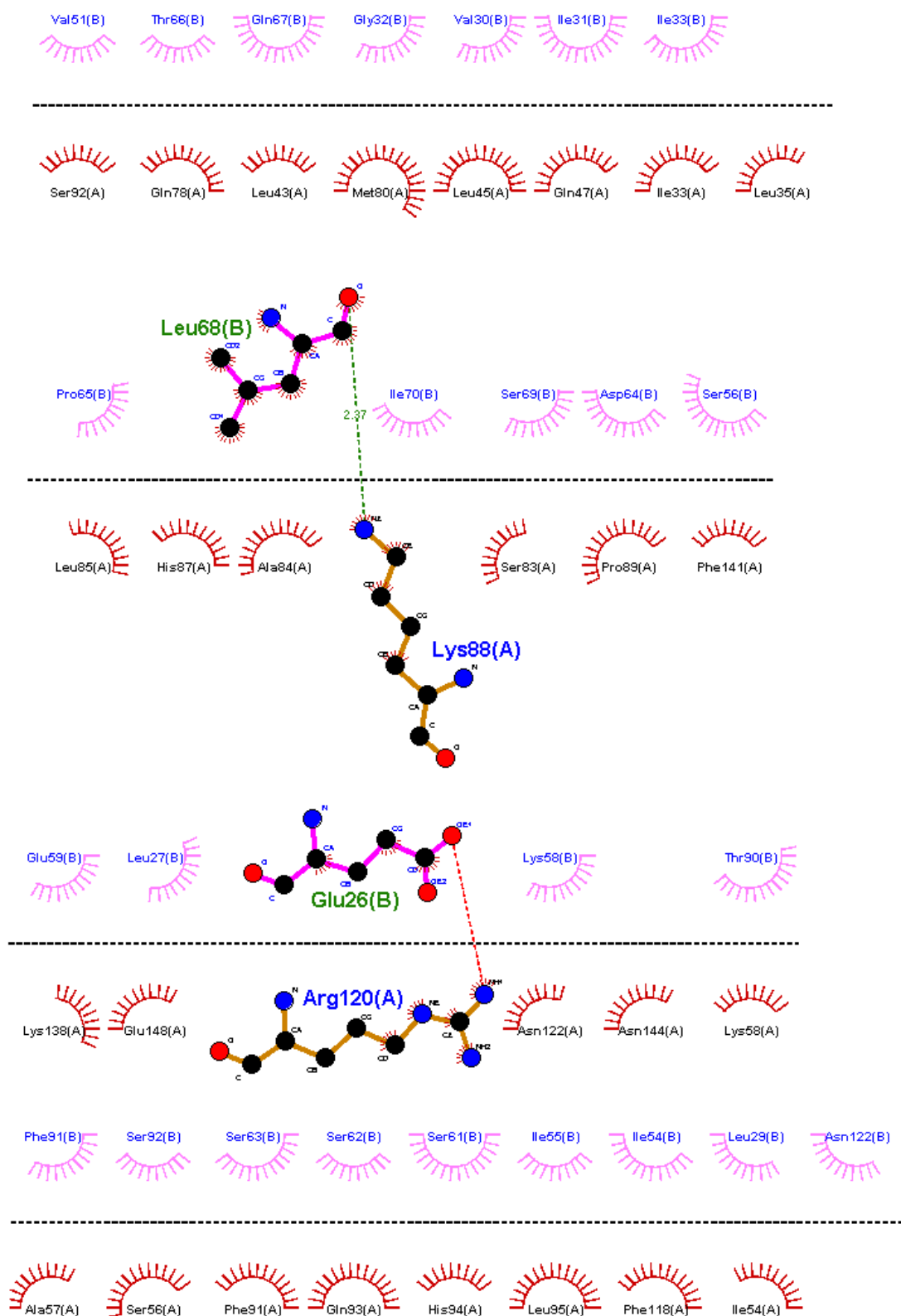
odpowiedzialnych za interakcje z błoną. Wynik dokowania monomerów przedstawiam na rysunku 28.



Rysunek 28: Model dimeru GerAC. Kolorem niebieskim oznaczony jest monomer A, natomiast kolorem różowym- monomer B. W obu monomerach brakuje 24 reszt aminokwasowych N-końca.

Monomery GerAC oddziałują ze sobą głównie w obrębie domeny I. Wiązanie wodorowe tworzy się między resztą Leu68 z monomeru B oraz Lys88 z monomeru A. LigPlot wskazuje również wiązanie kowalencyjne pomiędzy resztami Glu26(B) oraz Arg120(A), które, tak jak w przypadku innych oddziaływań między podjednostkami prawdopodobnie jest artefaktem oprogramowania. Być może w tym miejscu zachodzi inny rodzaj oddziaływań. Pozostałe interakcje to oddziaływania hydrofobowe (rysunek 29). W oddziaływaniach hydrofobowych udział bierze 25 reszt aminokwasowych monomeru A oraz 27 reszt monomeru B. Oddziaływania te są rozłożone mniej symetrycznie niż oddziaływania w dimerach GerAA i GerAB.

Dodatkowo, porównując model dimeru GerAC do dimeru homologa, łatwo zauważyć różnicę w umiejscowieniu domen I. W dimerze GerBC domeny te się wyraźnie przeplatają, natomiast w dimerze GerAC oddziałują ze sobą, ale są umiejscowione w przestrzeni między domeną I oraz II.



Rysunek 29: Schemat oddziaływań między monomerami A i B dimeru GerAC.

6.4.2.4. *Proponowany model receptora wykorzystujący dimery białek błonowych*

Wykorzystując dimery uzyskane za pomocą dokowania białko-białko, utworzyłam potencjalny model receptora GerA. Ponieważ białko C tego receptora jest jedynie zakotwiczone w błonie i znajduje się poza jej obszarem, w badaniach nad kompleksem receptora w pierwszej kolejności skupiłam się na oddziaływaniach między białkami błonowymi — dimerami GerAA oraz GerAB.

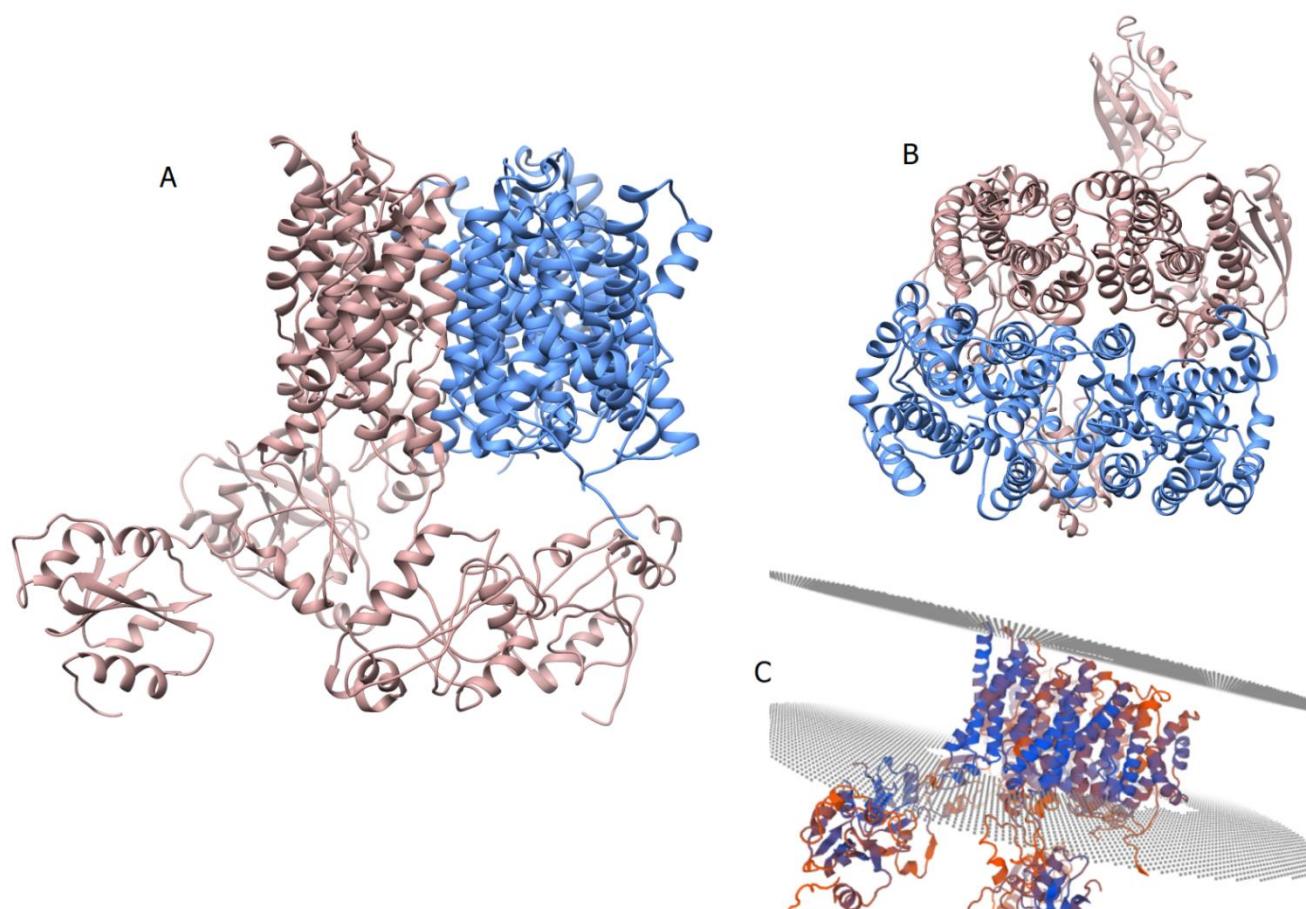
Kompleks błonowej części receptora, czyli dimerów GerAA-GerAB, zbudowałam z wykorzystaniem algorytmów dokowania zaimplementowanych w narzędziu ClusPro. Ze względu na swoją wielkość dokowanie kompleksu było bardzo kosztowne obliczeniowo, dlatego ClusPro — serwer działający na bostońskim superkomputerze — był najbardziej rozsądnym narzędziem do przeprowadzenia obliczeń. Kompleks został wybrany na podstawie swoich właściwości elektrostatycznych oraz hydrofobowych i jest pokazany na rysunku 30.

Dimer GerAB składa się z łańcuchów (monomerów) A i B, natomiast dimer GerAA z łańcuchów C i D. Łańcuch A dimeru GerAB bezpośrednio oddziałuje z łańcuchami C oraz D dimeru GerAA, natomiast łańcuch B dimeru GerAB oddziałuje tylko z łańcuchem D dimeru GerAA. Łańcuchy A (dimer GerAB) i C (dimer GerAA) łączą dwa wiązania wodorowe, wiązanie kowalencyjne oraz oddziaływania hydrofobowe. Wiązanie wodorowe formuje się między następującymi resztami aminokwasowymi: Leu338(A)-Gln434(B) oraz Tyr334(A)-Glu343(C). Reszty aminokwasowe Lys265(A)-Glu481(C) łączą się wiązaniem kowalencyjnym, co ponownie może być artefaktem oprogramowania obrazującego oddziaływania między białkami. Łańcuchy A i D łączą trzy oddziaływania hydrofobowe (Met409(D), Pro325(A) i Asp326(A)). Łańcuch B dimeru GerAB oraz łańcuch D dimeru GerAA połączone są dwoma wiązaniami wodorowymi: Met311(B)-Arg402(D) oraz Val163(B)-Gln305(D) i jednym wiązaniem kowalencyjnym Glu51(B)-Arg405(D), poza tym występują między nimi liczne oddziaływania hydrofobowe.

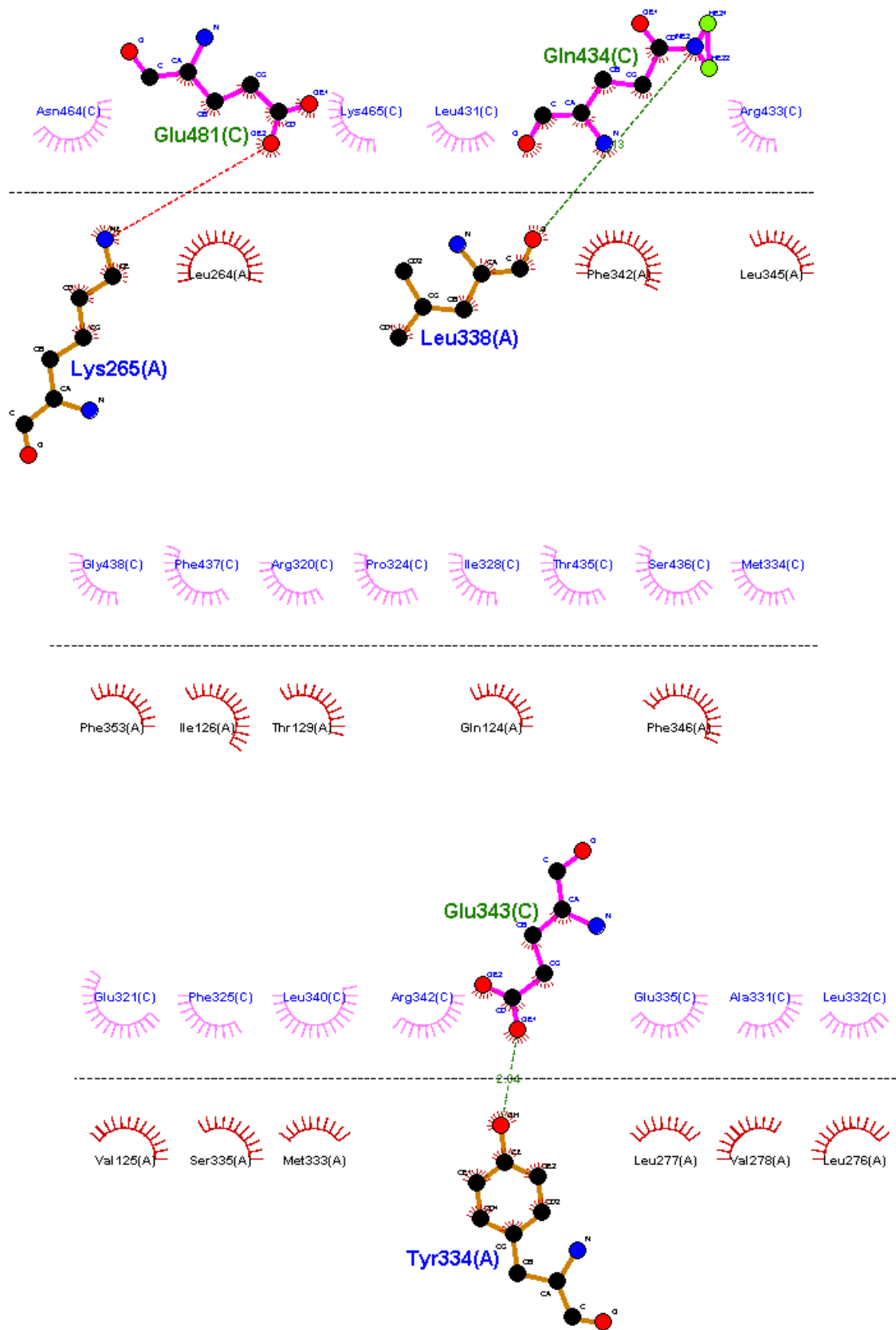
Przywoływane wcześniej badania dotyczące wpływów mutacji w białkach GerAA i GerAB na funkcjonowanie przetrwalników i inicjację procesu kiełkowania pokazują, że trzy z reszt aminokwasowych, które biorą udział we wiązaniu do siebie dimerów GerAA i GerAB w proponowanym kompleksie, są istotne dla prawidłowego funkcjonowania receptora. Kwas glutaminowy 51 białka GerAB nie jest co prawda resztą konserwowaną w innych podjednostkach B, jednak jej substytucja na leucynę sprawia, że przetrwalniki charakteryzują się nie działającym receptorem GerA. Co ciekawe, w tym przypadku

obecność kompleksu receptora jest wykrywalna (Cooper i Moir, 2011). Metionina 409 białka GerAA biorąca udział w oddziaływaniach hydrofobowych między łańcuchami A i D kompleksu, w eksperymentach została zastąpiona asparaginą, co w efekcie skutkowało bardzo niewielką lub zerową odpowiedzią receptora na obecność germinantów. Kwas glutaminowy 335 białka GerAA, wspomniany już w tej pracy jako reszta wysoce konserwowana wśród podjednostek A, znajduje się tuż obok motywu PFPP i wspólnie z resztą 330 wskazywana jest przez badania jako istotna dla poprawnego działania receptora ze względów strukturalnych. Substytucja E335A skutkuje brakiem odpowiedzi receptora na obecność L-alaniny w środowisku oraz utratą 95% ilości białka GerAC (Mongkoltharuk, Cooper, Mawer, Allan i Moir, 2011).

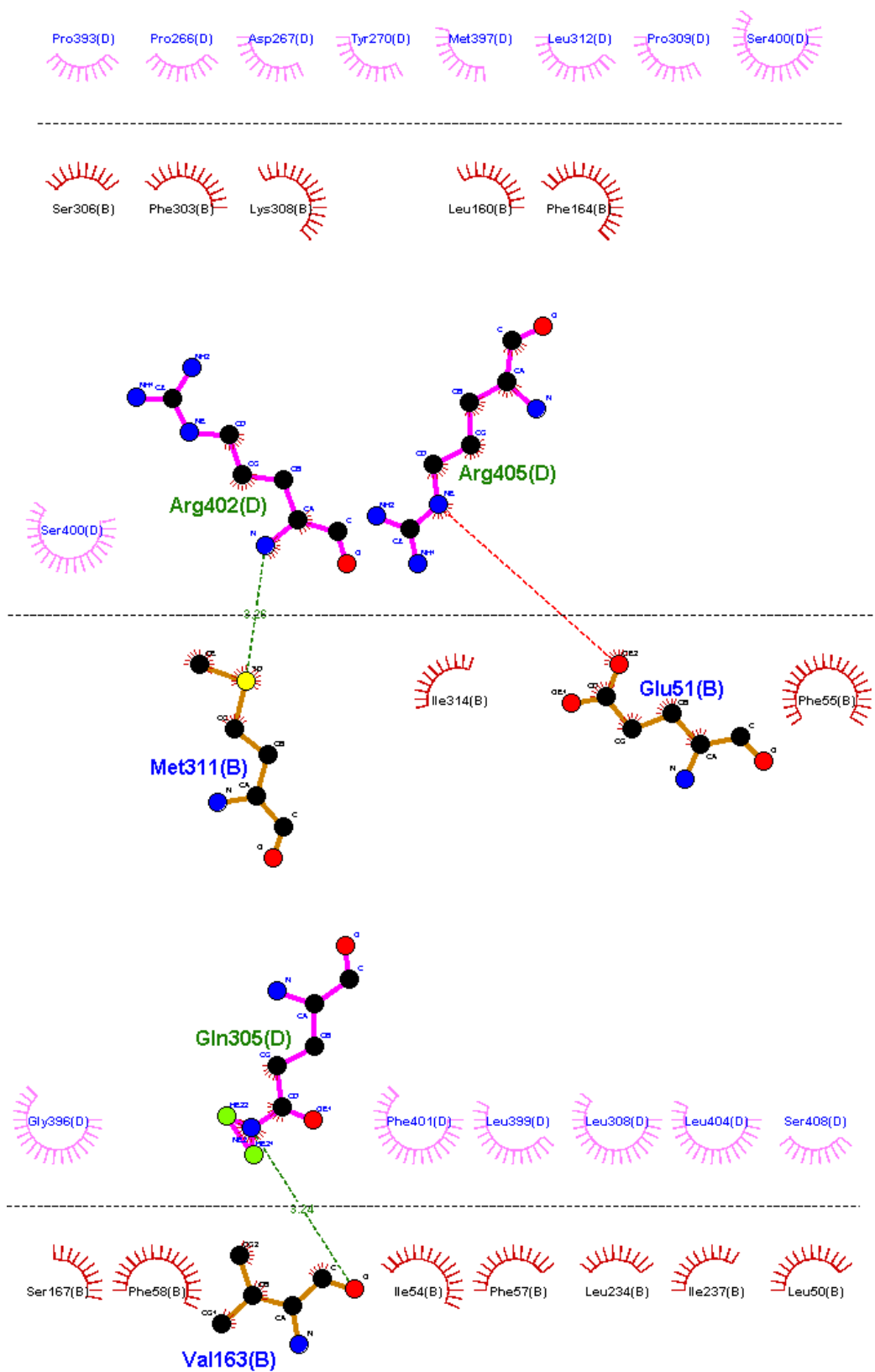
Wyniki oceny modelu kompleksu na podstawie QMbrane są bardzo zbliżone do wyników, jakie osiągały monomery GerAA i GerAB. *Normal score* wynosi 0,66, natomiast *z-score* jest równy -0,5. Dodatkowo, 86,52% reszt aminokwasowych znajduje się w obszarze uznawanym za dozwolony na wykresie Ramachandrana.



Rysunek 30: Błonowa część modelu receptora GerA. A) Kompleks dimerów GerAA i dimerów GerAB. Kolorem niebieskim zaznaczony jest dimer podjednostki B, natomiast kolorem różowym – dimer podjednostki A. B) Kompleks widziany od góry. C) Potencjalna orientacja kompleksu w błonie obliczona za pomocą algorytmu QMbrane.



Rysunek 31: Oddziaływania między łańcuchami A i C kompleksu receptora zbudowanego z dimerów GerAA i GerAB.

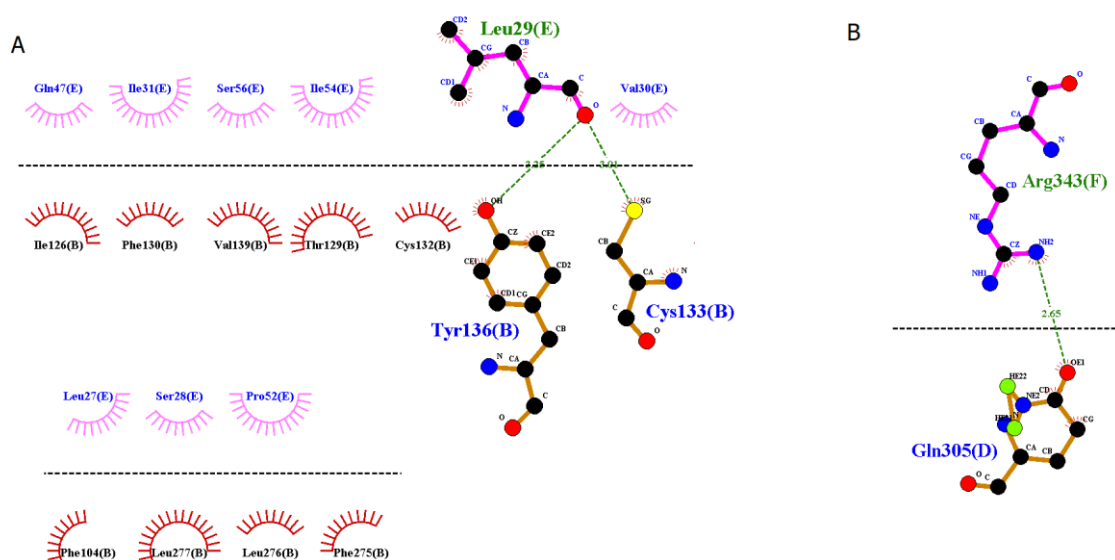


Rysunek 32: Oddziaływania między łańcuchami B i D kompleksu receptora GerA zbudowanego w oparciu o dimery GerAA i GerAB.

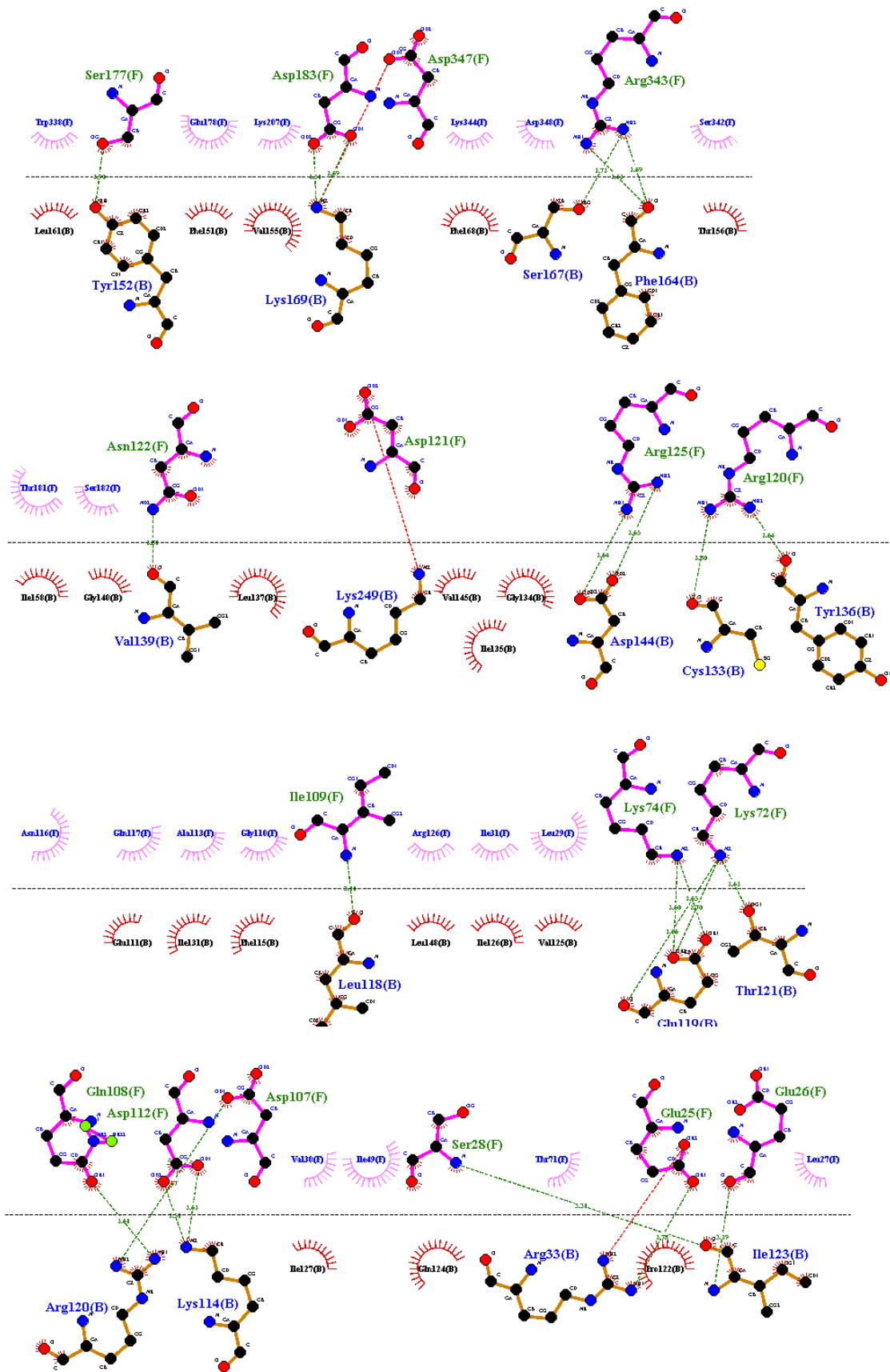
6.4.2.5. Kompleks GerA zbudowany na podstawie dimerów podjednostek

Model receptora GerA (rysunek 35) również zbudowałam przy użyciu serwera ClusPro, a energia potrzebna do zadokowania dimeru GerAC do kompleksu dimerów GerAA i GerAB wyniosła -1947,2 kJ/mol. Widoczną wadą tego modelu jest to, że część dimeru GerAC, według oceny QMbrane, jest zanurzona w błonie. Trudności z dokowaniem GerAC mogą wynikać z braku pętli na N-końcu monomerów tego dimeru, która została usunięta w celu zbudowania dimerów. Właśnie ta pętla jest prawdopodobnie odpowiedzialna za zakotwiczenie lipoproteiny GerAC w błonie.

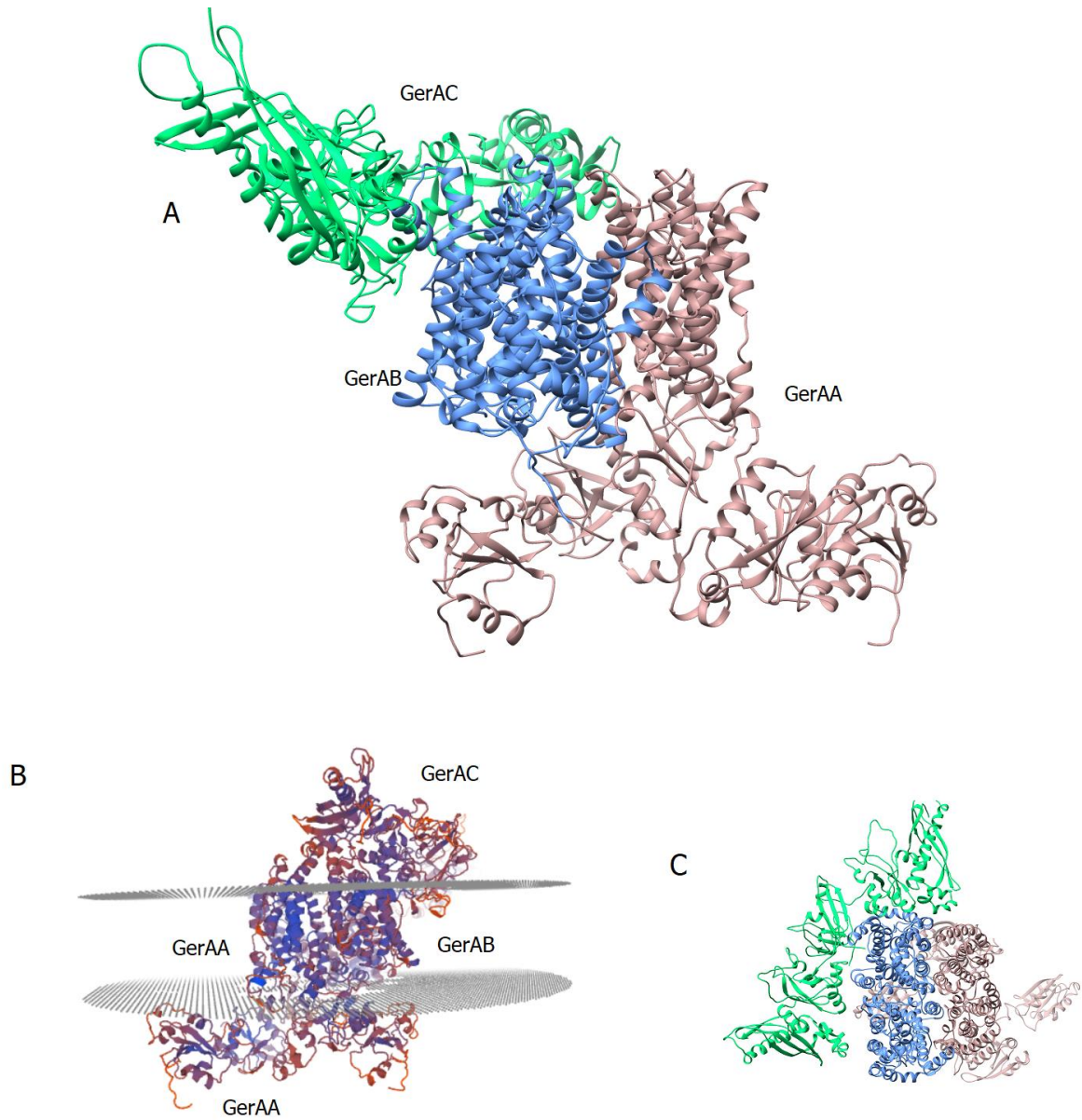
Dimer GerAC (łańcuchy E i F) w tym modelu receptora oddziałuje z oboma monomerami GerAB (łańcuchy A i B) oraz z jednym z monomerów GerAA (łańcuchem D). Łańcuchy D i F oddziałują ze sobą za pomocą jedynie jednego wiązania wodorowego, natomiast łańcuchy B i E za pomocą dwóch wiązań wodorowych i oddziaływań hydrofobowych. Łańcuch F dimeru GerAC wiąże się do łańcucha B dimeru GerAB za pomocą 24 wiązań wodorowych, 3 wiązań kowalencyjnych (kolor czerwony na rysunku 34) oraz oddziaływań hydrofobowych.



Rysunek 33: Rozłożenie oddziaływań w kompleksie receptora GerA. A) Opis oddziaływań pomiędzy łańcuchem E dimeru GerAC, a łańcuchem B dimeru GerAB. B) Wiązanie wodorowe pomiędzy łańcuchem F dimeru GerAC a łańcuchem D dimeru GerAA.



Rysunek 34: Rozłożenie oddziaływań pomiędzy łańcuchem B dimeru GerAB, a łańcuchem F dimeru GerAC.

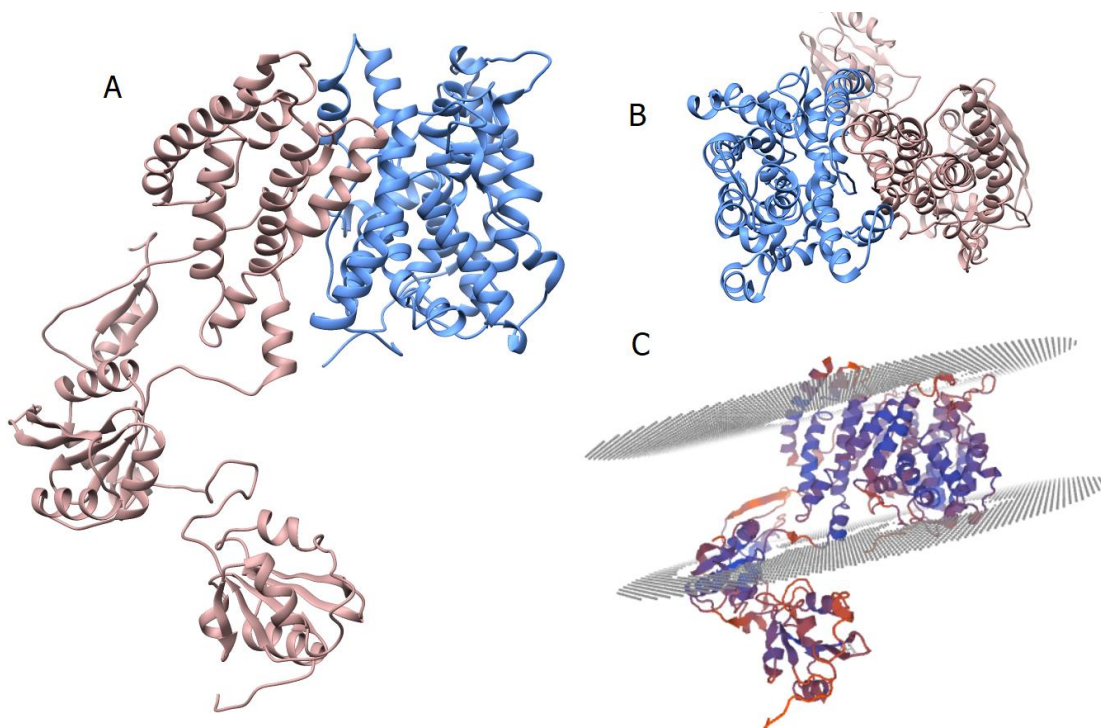


Rysunek 35: Model kompleksu GerA zbudowany na podstawie dimerów podjednostek. A) Model przedstawiający budowę receptora GerA. Kolorem zielonym zaznaczono dimer GerAC, kolorem niebieskim dimer GerAB, natomiast kolorem różowym – dimer GerAA. B) Potencjalne ułożenie receptora w błonie. C) model receptora widziany z góry.

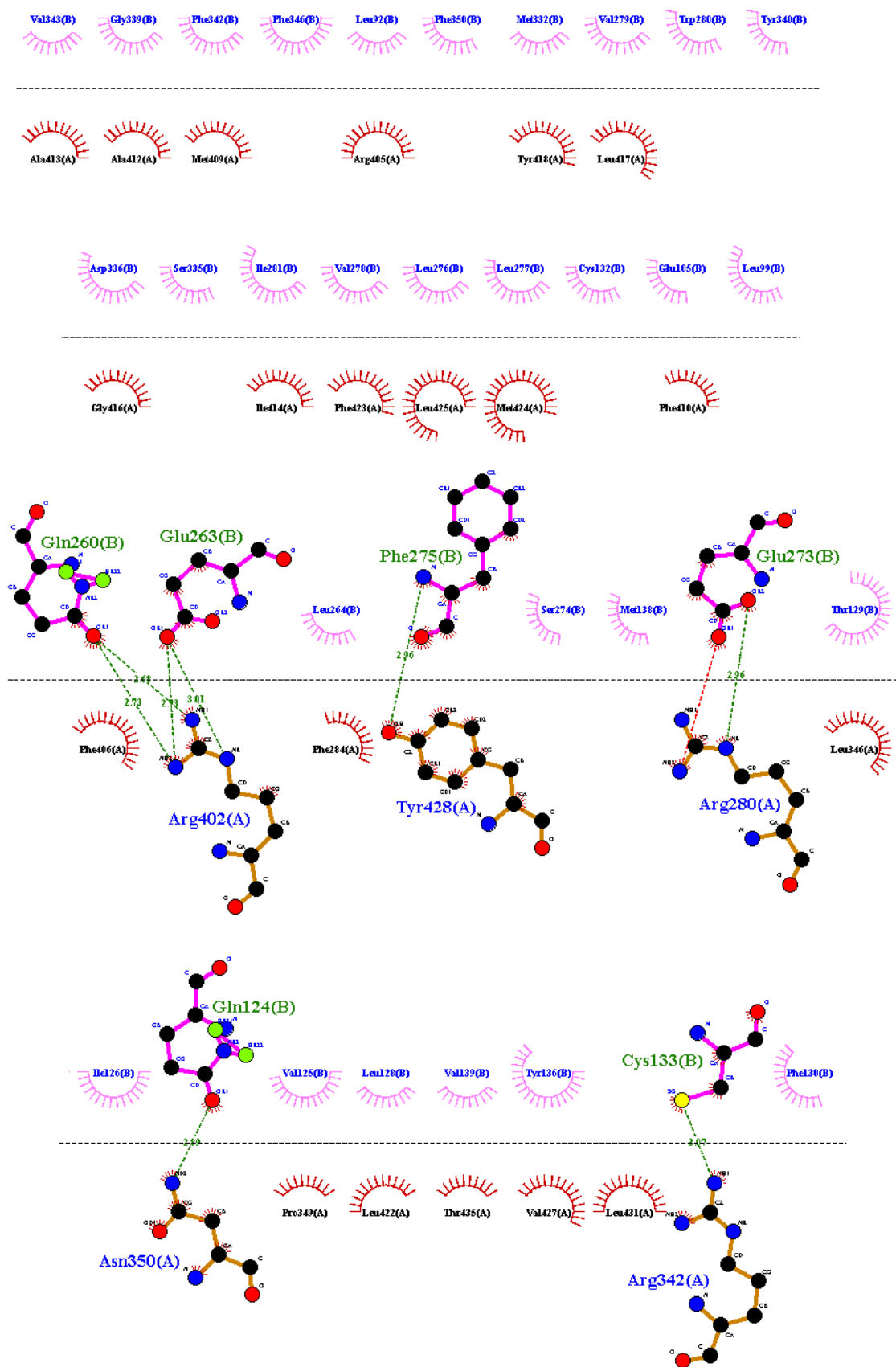
6.4.3. Kompleks monomerów białek błonowych receptora GerA

Błonowa część receptora GerA zbudowana z monomerów GerAA i GerAB powstała przy wykorzystaniu serwera ClusPro. Wyniki uzyskane z dokowania białko-białko za pomocą tego narzędzia dały kilka interesujących kompleksów. Do dalszych badań wybrałam kompleks, który był najbardziej faworyzowany ze względów hydrofobowych — w klastrze opisującym taki model znalazły się aż 74 kompleksy, a energia dokowania wynosiła -2710,6 kJ/mol. Reszty pozytywnie oceniane przez wykres Ramachandrana wynoszą 85,68%.

Wstępne obliczenia orientacji kompleksu w błonie pokazują, że fragment 238-264 białka GerAA oraz reszty 177-193, wchodzące w skład domeny niebłonowej białka mogą być zanurzone w błonie. Fragment 238-264 został przeze mnie przewidziany jako helisa błonowa przez algorytmy do przewidywań regionów transbłonowych (tabela 3). Powstały kompleks stabilizowany jest przez sześć wiązań wodorowych oraz liczne oddziaływania hydrofobowe. Monomer GerAA oznaczony jest jako łańcuch A, natomiast monomer GerAB jako łańcuch B. Wiązania wodorowe tworzą się pomiędzy następującymi parami reszt aminokwasowych: Gln260B-Arg402A, Gln263C-Arg402A, Phe275B-Tyr428A, Gln273B-Arg280A, Gln124B-Asn350A oraz Cys133B-Arg342A.



Rysunek 36: Model kompleksu błonowej części receptora oparty na monomerach podjednostek A i B. A) Model kompleksu – widać delikatne nachylenie podjednostki A względem podjednostki B. B) Widok z góry na kompleks GerAA-GerAB. C) Wstępna orientacja kompleksu w błonie. Widzimy, że błona rozciąga się na część niebłonowej domeny GerAA.

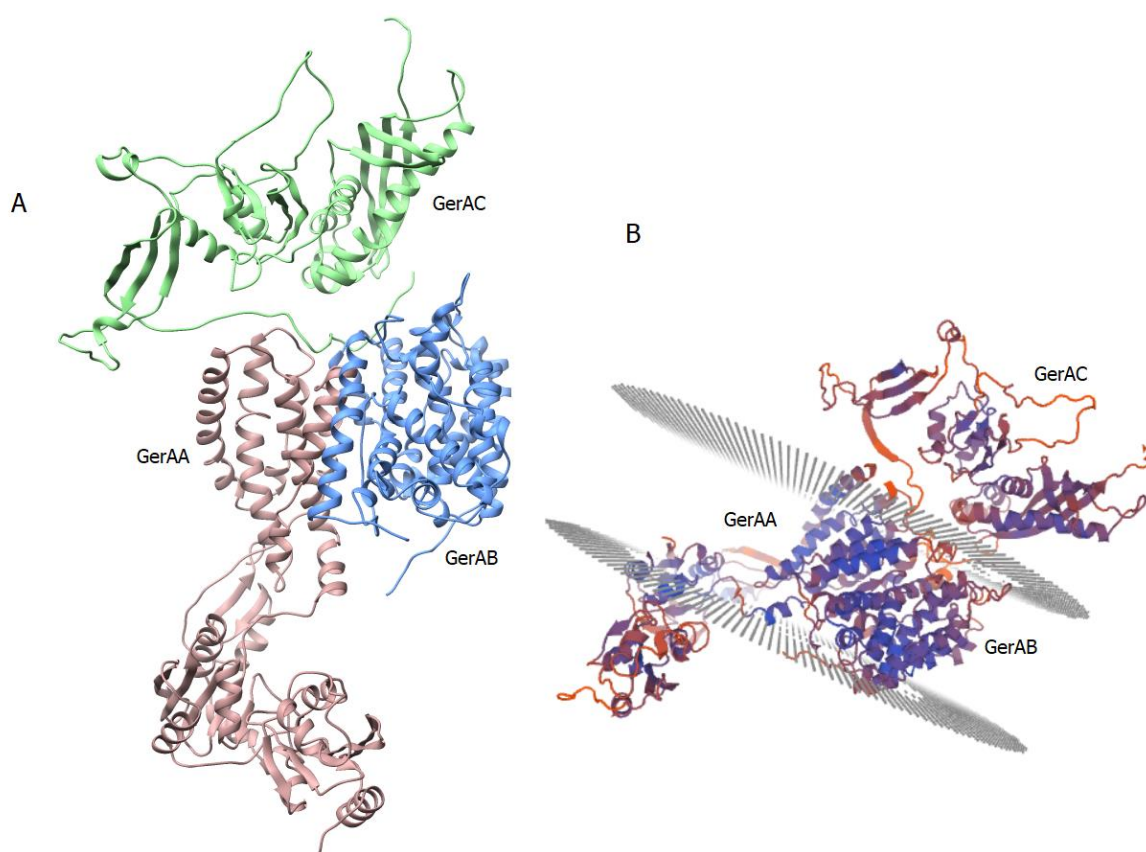


Rysunek 37: Interakcje między monomerami GerAA (łańcuch A) oraz GerAB (łańcuch B) w kompleksie.

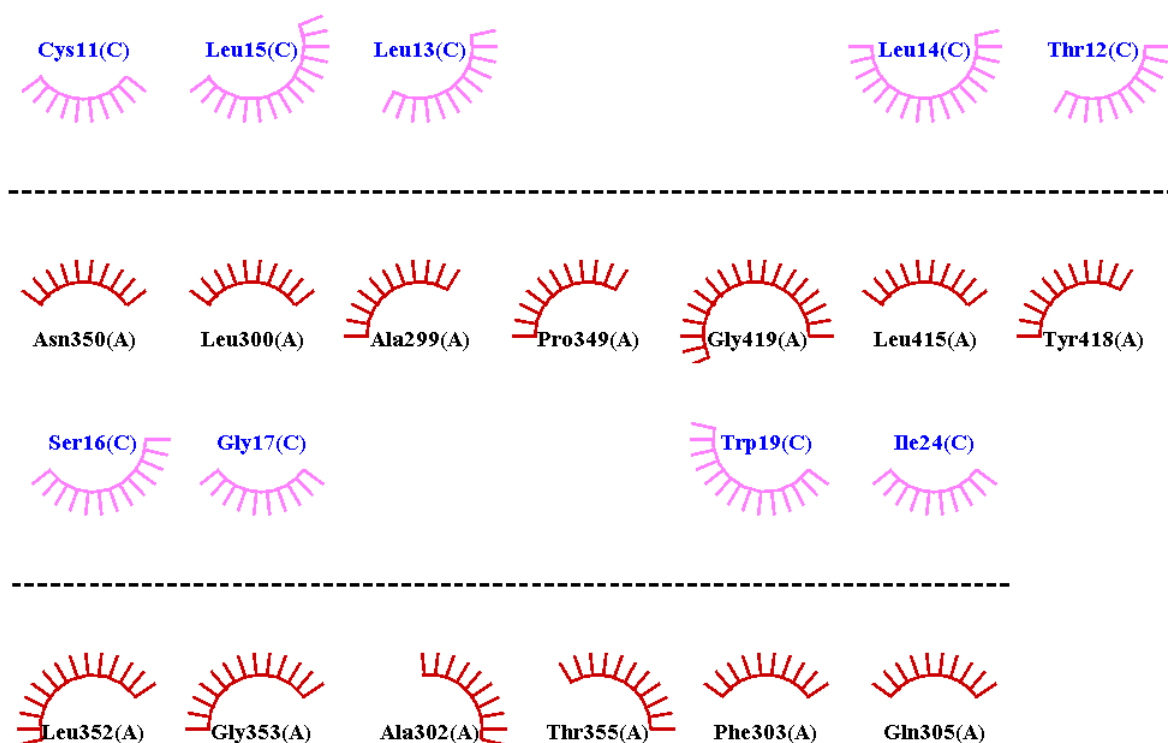
6.4.3.1. Receptor GerA zbudowany na podstawie monomerów

Podjednostka C receptora GerA zadokowałam do kompleksu zbudowanego z GerAA i GerAB za pomocą ClusPro. Serwer ten dał dwa ciekawe wyniki — pierwszy z nich znajduje się na rysunku 36 i pokazuje, że podjednostka C znajdować się będzie w bardziej zewnętrznej części przetrwalnika. Kompleks ten osiągnął energię równą $-1840,8$ kJ/mol, a ułożenie podjednostek jest zgodne ze schematem rozmieszczenia białek receptora GerA pokazanym na rysunku 2 i szeroko stosowanym w literaturze.

Drugi interesujący wynik (niepokazany w pracy) zakładał, że białko GerAC wiąże się z niebłonową domeną białka GerAA, a tym samym znajduje się w rdzeniu przetrwalnika. Ten kompleks dokował się do siebie z energią równą $-1405,9$ kJ/mol.



Rysunek 38: Receptor GerA zbudowany z monomerów podjednostek. A) Struktura receptora: kolorem zielonym zaznaczona jest podjednostka C receptora, która znajduje się po zewnętrznej stronie wewnętrznej błony przetrwalnika, kolorem niebieskim podjednostka B, a kolorem różowym – podjednostka A, która znajduje się zarówno w błonie jak i w rdzeniu przetrwalnika B) Schematyczna orientacja kompleksu w błonie obliczona za pomocą narzędzia QMBrane.

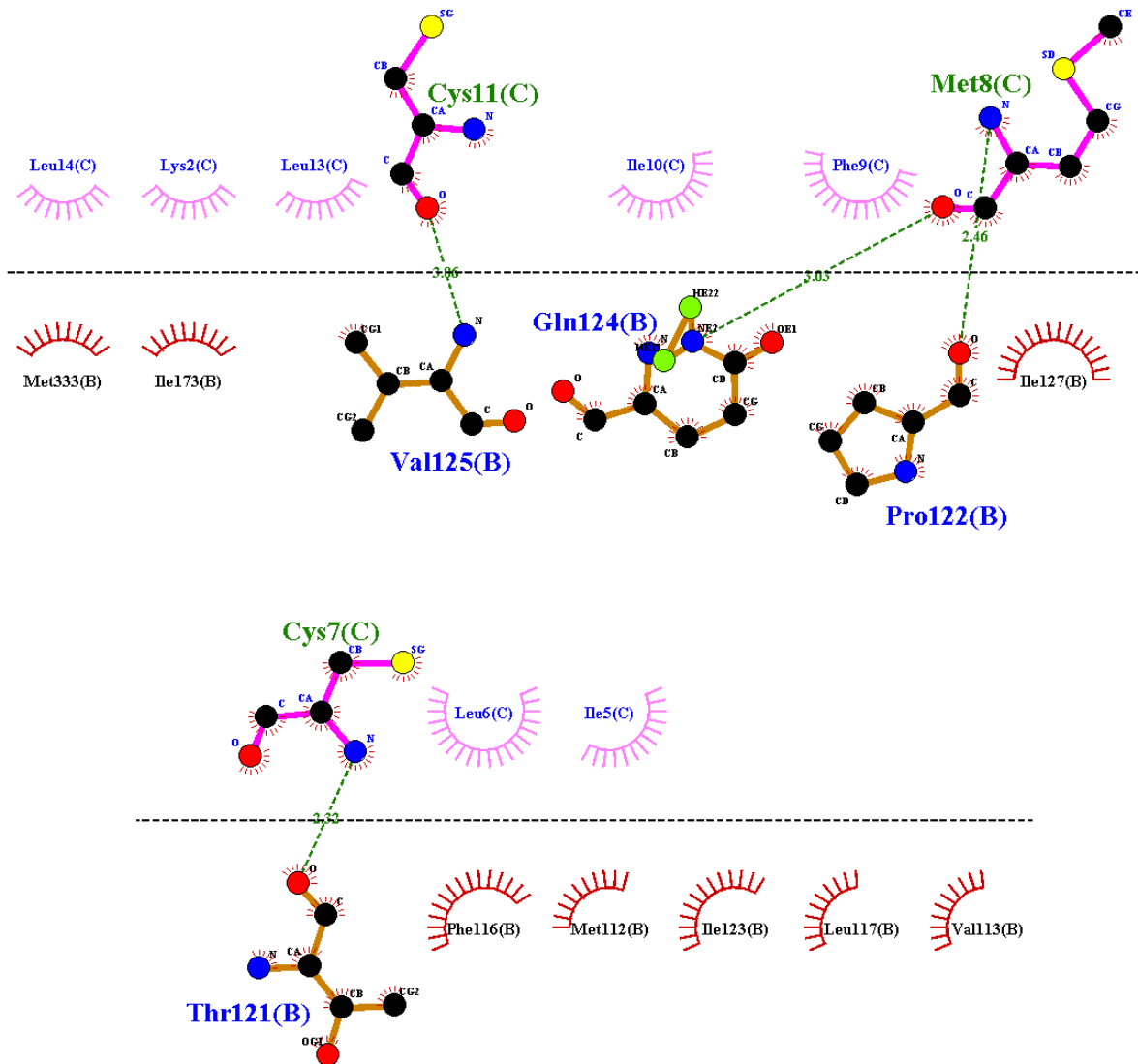


Rysunek 39: Schemat oddziaływań pomiędzy białkami A i C w kompleksie receptora GerA.

W receptorze GerA zbudowanym z monomerów podjednostek, GerAC oddziałuje zarówno z podjednostką A jak i z podjednostką B. Przy czym, oddziaływania pomiędzy podjednostkami A i B mają charakter hydrofobowy (rysunek 36), natomiast oddziaływania białka C z GerAB stabilizowane są czterema wiązaniami wodorowymi (rysunek 37). Warty zaznaczenia jest, że w oddziaływaniach pomiędzy A i C udział biorą reszty 299 oraz 302 białka GerAA, które, jak wykazała dr Anna Grela, są istotne dla poprawnego funkcjonowania receptora. W oddziaływaniach pojawia się również reszta 300, która wraz z resztą 299 często pojawiają się w uzyskanych przeze mnie wynikach dokowania germinantów do białka GerA, jego dimerów oraz kompleksów GerA.

Oddziaływania ze strony GerAC są skupione, zarówno w przypadku oddziaływania z GerAA jak i z GerAB, w regionie N-terminalnym tego białka, dokładniej w długiej nieustrukturyzowanej pętli, której część według struktury drugorzędowej powinna przyjąć kształt α -helisy. Jest to region, w którym znajduje się reszta Cys18 odpowiedzialna za połączenie z wewnętrzną błoną kotwicą diacyloglicerolową.

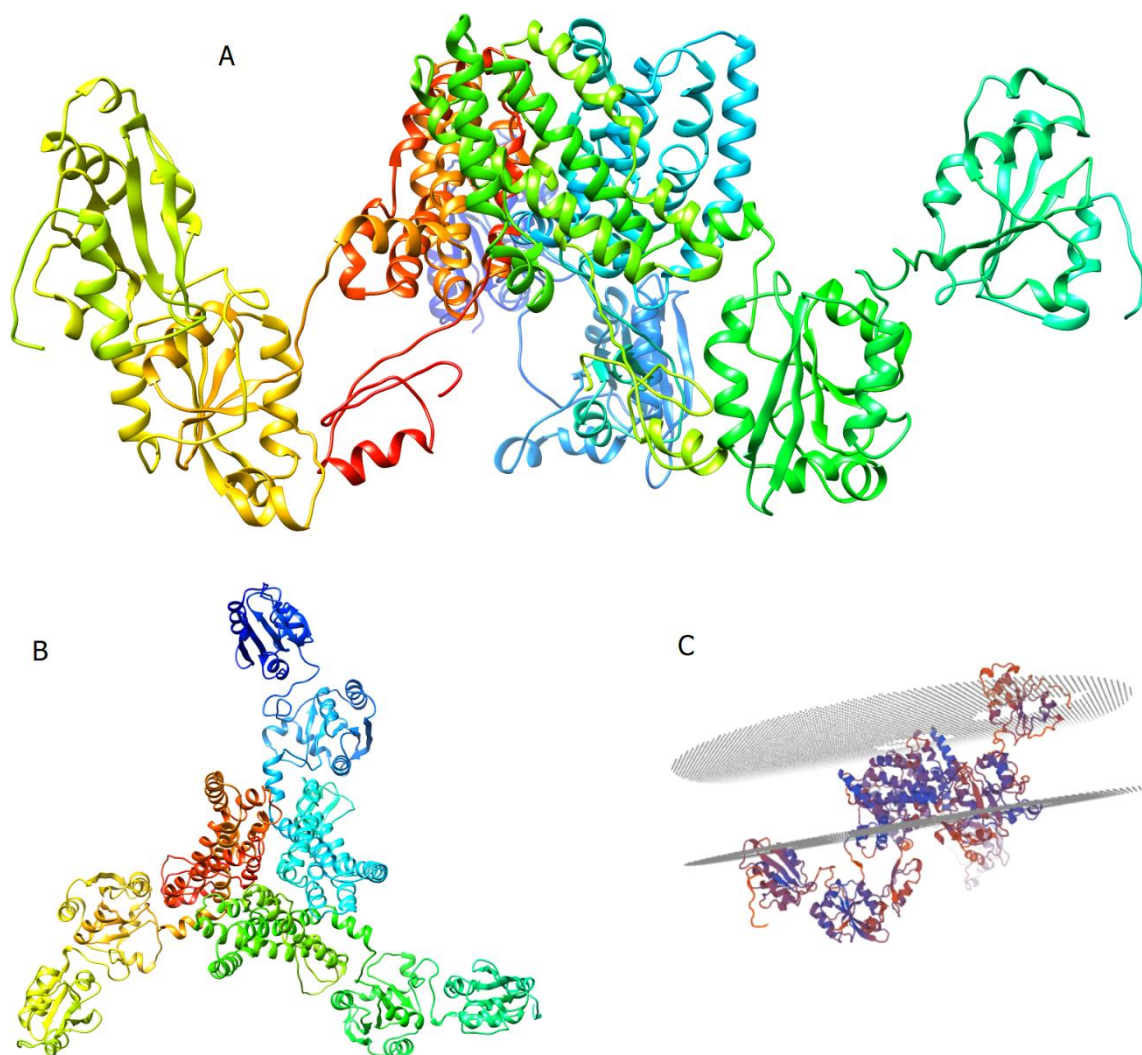
GerAC oddziałuje z GerAB za pomocą oddziaływań hydrofobowych oraz następujących wiązań wodorowych: Cys11C-Val125B, Met8C-Gln124B, Met8C-Pro122B oraz Cys7C-Thr121B.



Rysunek 40: Schemat oddziaływań pomiędzy białkami GerAB oraz GerAC w kompleksie receptora GerA.

6.4.4. Trimery podjednostek

Trimer podjednostki GerAA zbudowałam przy wykorzystaniu serwera ClusPro. Energia niezbędna do zadokowania monomerów w kompleks wynosi $-150,3$ kJ/mol, a wybrany kompleks charakteryzuje się tym, że najlepiej wypadł podczas zaimplementowanej w ClusPro oceny hydrofobowości.

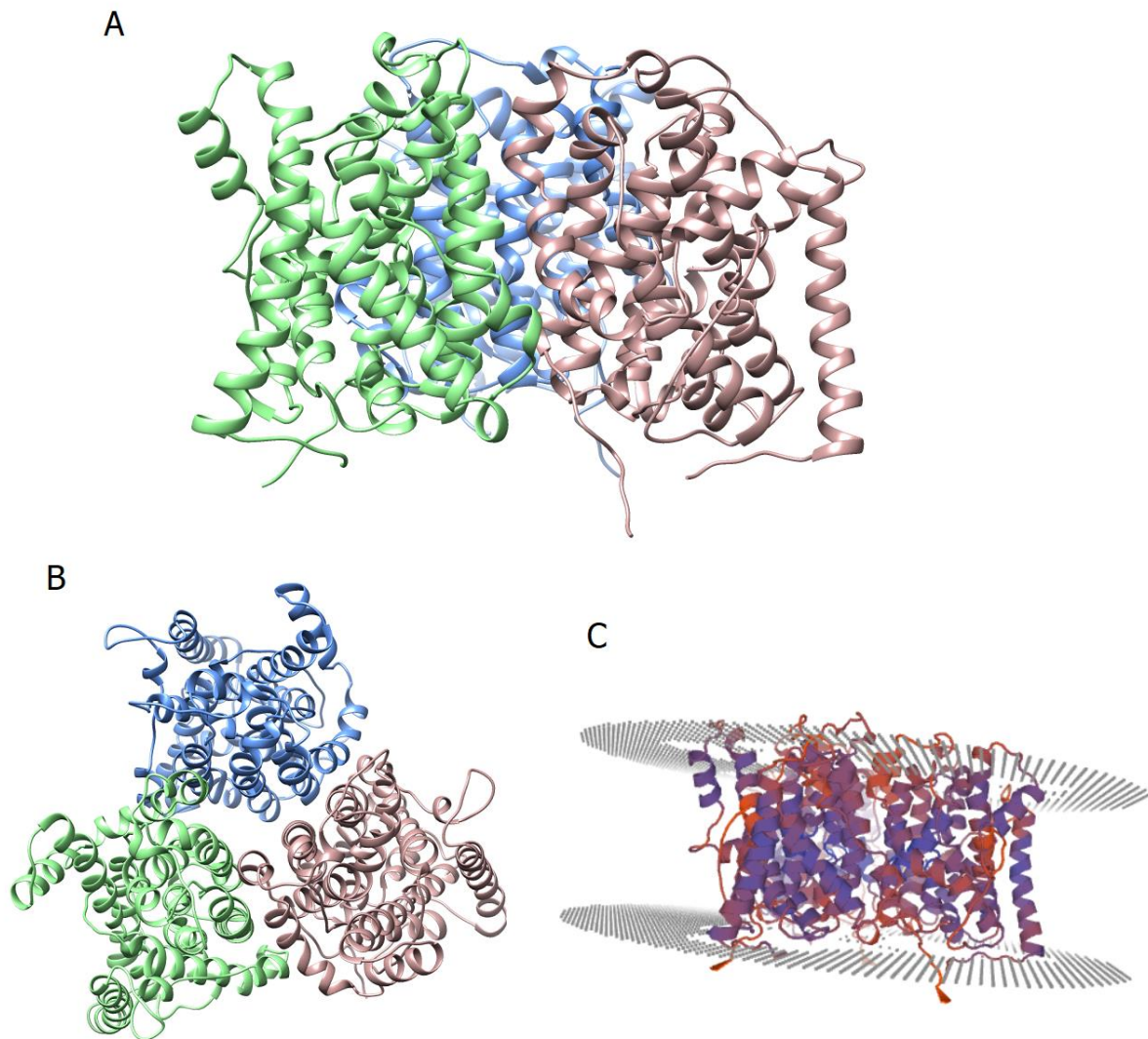


Rysunek 41: Model trimeru GerAA. A) Model kompleksu trimerów GerAA B) Rzut z góry na kompleks trimerów. C) Orientacja w błonie trimeru. Błona oznaczona jest kolorem szarym.

Rysunek 35 przedstawia model trimeru GerAA oraz obliczoną za pomocą serwera QMbrane orientację tego kompleksu w błonie. Jak widać, monomery dokują się do siebie pod pewnym kątem, co potem rzutuje na umiejscowienie kompleksu w błonie. Na części C

rysunku widzimy, że jeden z monomerów praktycznie w całości — łącznie z domeną niebłonową znajduje się w błonie.

Trimer GerAB powstał poprzez wykorzystanie dokowania symetrycznego w symetrii C3 i oprogramowania Sam. Energia wiązania monomerów w kompleks wyniosła - 513,75 kJ/mol. Trimery nie zostały przez mnie wykorzystane do dalszych badań (dokowania germinantów, budowa kompleksu receptora), stanowią natomiast istotne spojrzenie na to, jak może wyglądać kompleks receptora GerA.



Rysunek 42: Model trimery GerAB. A) Struktura kompleksu trimery. B) Widok na kompleks z góry. C) Orientacja trimery w błonie fosfolipidowej.

6.5. Dokowanie germinantów

Dostępne dane literaturowe podają, że germinanty (L-alanina i L-walina) inicjujące proces kiełkowania przetrwalników *B. subtilis*, prawdopodobnie oddziałują z podjednostką A i/lub B receptora GerA. Jak dotąd w literaturze nie pojawiły się przesłanki stojące za tym, by podjednostka C mogła być bezpośrednio zaangażowana w wiązanie germinantów do receptora.

Wykorzystując metody dokowania molekularnego oraz modele struktur trzeciorzędowych GerAA (w pracy zostały opisane tylko dokowania do drugiego modelu białka A, wyniki dla modelu pierwszego zostały pominięte) i GerAB, wytypowałam regiony, które mogą odpowiadać za interakcje tych białek z germinantami. Dokowania zostały przeprowadzone zarówno dla monomerów podjednostek jak i dimerów oraz dla modeli kompleksu błonowej części receptora.

Ponieważ do niedawna w literaturze można było znaleźć bardzo niewiele informacji o kieszeniach wiążących receptora GerA, głównym krokiem tej części eksperymentu było dokowanie globalne germinantów do podjednostek. Dokowanie globalne pozwoliło mi wytypować potencjalne miejsca wiązania na podstawie częstotliwości wiązania się liganda do danego regionu białka. Wszystkie dokowania germinantów przeprowadziłam z użyciem programu AutoDock Vina, co dokładniej opisałam w rozdziale 5.7.2 *Dokowanie ligand-receptor*.

6.5.1. Dokowania germinantów do podjednostki A

6.5.1.1. Wyniki przewidywań miejsc oddziaływań z ligandem dla GerAA

Serwer PLIP typuje cztery regiony białka GerAA, które mogą być miejscami oddziaływania tego białka z ligandami. Pierwsza przewidywana przez serwer kieszeń wiążąca obejmuje reszty Leu14 oraz Ala15. Drugi region to reszty Arg151, Ser155 i Leu159. Trzeci: Leu307 i Leu308; czwarty region: Ile338, Val426, Arg433 oraz Gln434. Miejsca te znajdują się poza regionami transbłonowymi. Trzeci i czwarty region to miejsca znajdujące się w bardziej zewnętrznej części przetrwalnika, natomiast regiony pierwszy i drugi to fragmenty dużej domeny pozabłonowej, umiejscowionej w wewnętrznej warstwie spory (rdzeniu).

Serwer BiteNet również wskazuje cztery potencjalne kieszenie wiążące ligandy w białku GerAA. Są to kolejno: 1) Asn13, Leu16, Lys20, Lys89, Thr91, Ile92; 2) Ser281, Arg283, Phe284, Ale285, Ile287; 3) Arg283, Ala285, Ser286, Ile287, Ile289, Thr290, Leu378, Val379, Gly398, Leu399, Phe401, Val403 oraz 4) Phe406, Ile407, Phe410, Phe423, Leu425.

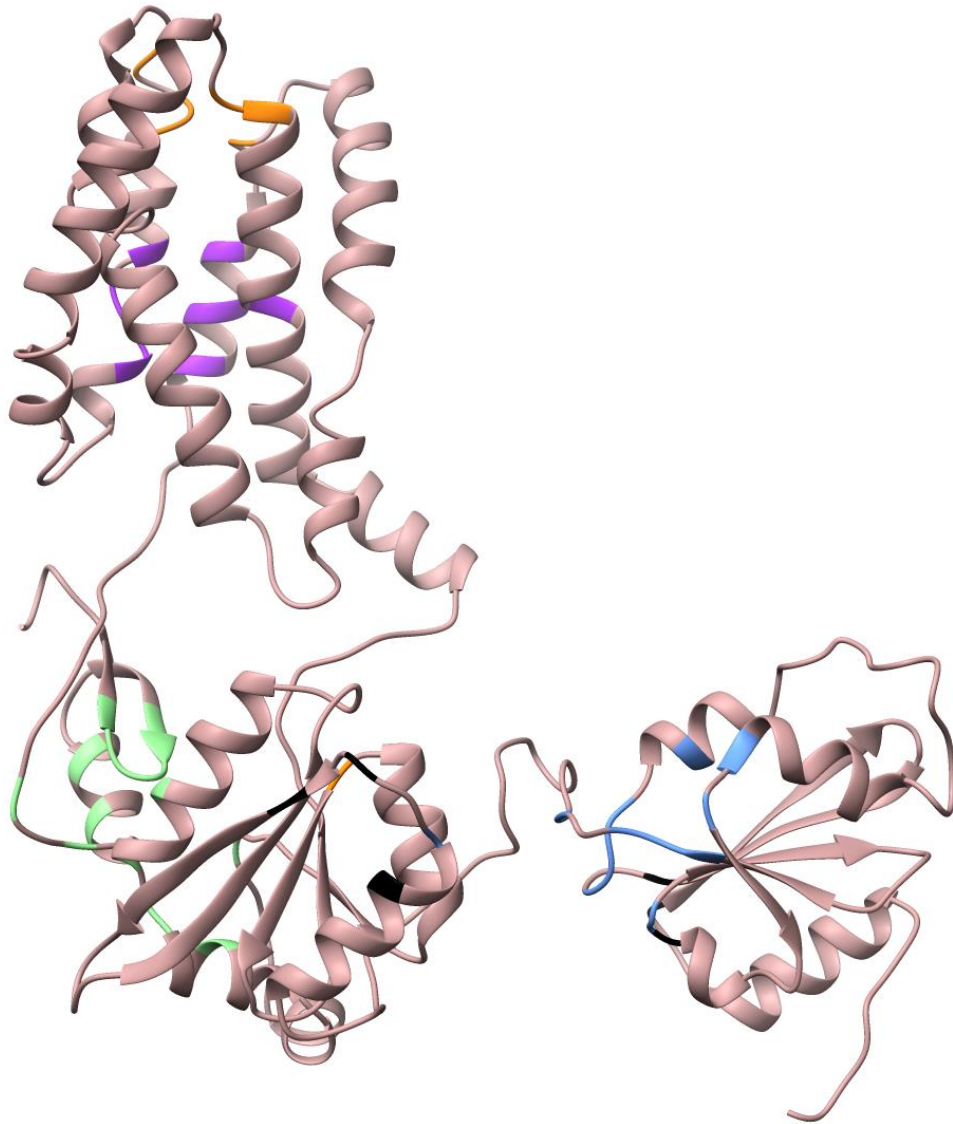
6.5.1.2. Wyniki dokowania globalnego germinantów do monomeru GerAA

Wyniki dokowań globalnych L-alaniny oraz L-waliny do monomeru GerAA zebrałam w tabeli 7. Dokowanie globalne L-alaniny określa dziewięć potencjalnych kieszeni wiążących, natomiast dokowanie L-waliny wskazuje siedem regionów. Niektóre z reszt przewidzianych przez serwery wykorzystane do wstępnego poszukiwania miejsc wiążących (opisane w punkcie wyżej) również pojawiają się wśród reszt aminokwasowych biorących udział w interakcjach z germinantami podczas dokowań globalnych.

Klastry, czyli zbiory kompleksów o podobnym ułożeniu liganda w receptorze, dla L-waliny oraz L-alaniny są bardzo zbliżone do siebie pod względem lokalizacji w białku. Trzy z nich wydają się szczególnie interesujące ze względu na liczbę kompleksów oraz na swoje położenie. Klastry numer 2 dla L-alaniny i L-waliny są najliczniej reprezentowane (kolor zielony na rysunku 43). Region ten znajduje się w części niebłonowej białka, na styku ogona C-końcowego białka oraz drugiej domeny części niebłonowej. Klaster 5 L-alaniny oraz 1 L-waliny umiejscowiony jest w pierwszej domenie niebłonowej białka (kolor niebieski). Trzeci wspólny region — klastry numer 4 obu germinantów — to region obejmujący fragment helisy opisywanej w podrozdziale Pozycje 299 i 302. Reszty 299 oraz 300 często wskazywane były jako reszty aminokwasowe tworzące wiązania wodorowe z germinantami. Reszta 299, jak pokazują nasze badania (Grela i inni, 2018), jest istotna dla poprawnego funkcjonowania receptora. Region ten na rysunku 43 zaznaczony jest kolorem pomarańczowym — znajduje się on blisko zewnętrznej krawędzi błony i częściowo wystaje poza błonę, co czyni go regionem o wysokiej dostępności dla liganda. Dodatkowo na strukturze GerAA zazaczyłam, kolorem fioletowym, pierwszy klaster typowany w dokowaniach L-alaniny. Jest to region w części błonowej białka, a występujące w nim reszty aminokwasowe odpowiadają przewidywaniom serwera BiteNet.

Publikacja opisująca strukturę niebłonowej części homologa GerAA wskazuje styk obu domen tej struktury jako miejsce wiązania glukozy oraz L-waliny (Li i inni, 2019). Reszty aminokwasowe odpowiedzialne za interakcje homologa z ligandami to: Asp50,

Ile71, Met130, Arg185, Pro191 oraz Trp194. Zestawienie sekwencji homologa oraz białka GerAA pokazuje, że odpowiadające tej kieszeni wiążącej reszty w białku GerAA są następujące: Asp26, Phe45, Leu96, Arg151, Pro157, Ile160. Klastry 6,7 oraz 1 dla L-waliny znajdują się bardzo blisko reszt odpowiadających regionowi, który w homologu uznawany jest za potencjalną kieszeń wiążącą. Na rysunku 43 te reszty aminokwasowe zostały oznaczone kolorem czarnym. Jak widać, umiejscowione są one zarówno w pierwszej, jak i drugiej domenie niebłonowej białka. Sugerować to może, że germinanty wiążą się właśnie w tych rejonach białka, a co za tym idzie — muszą przemieścić się przez całą szerokość błony, aby dostać się do kieszeni wiążącej, ponieważ region ten znajduje się poza błoną w rdzeniu spory.



Rysunek 43: Potencjalne kieszenie wiążące białka GerAA naniesione na model struktury białka. Kolorem czarnym zaznaczono reszty aminokwasowe, które odpowiadają kieszeni wiążącej homologa białka GerAA.

Tabela 7: Wyniki dokowań globalnych L-alaniny oraz L-waliny do modelu białka GerAA. Zestawienie wyników zawiera liczbę kompleksów w poszczególnych klastrach oraz reszty aminokwasowe białka wskazywane jako te, które oddziaływały z germinantami.

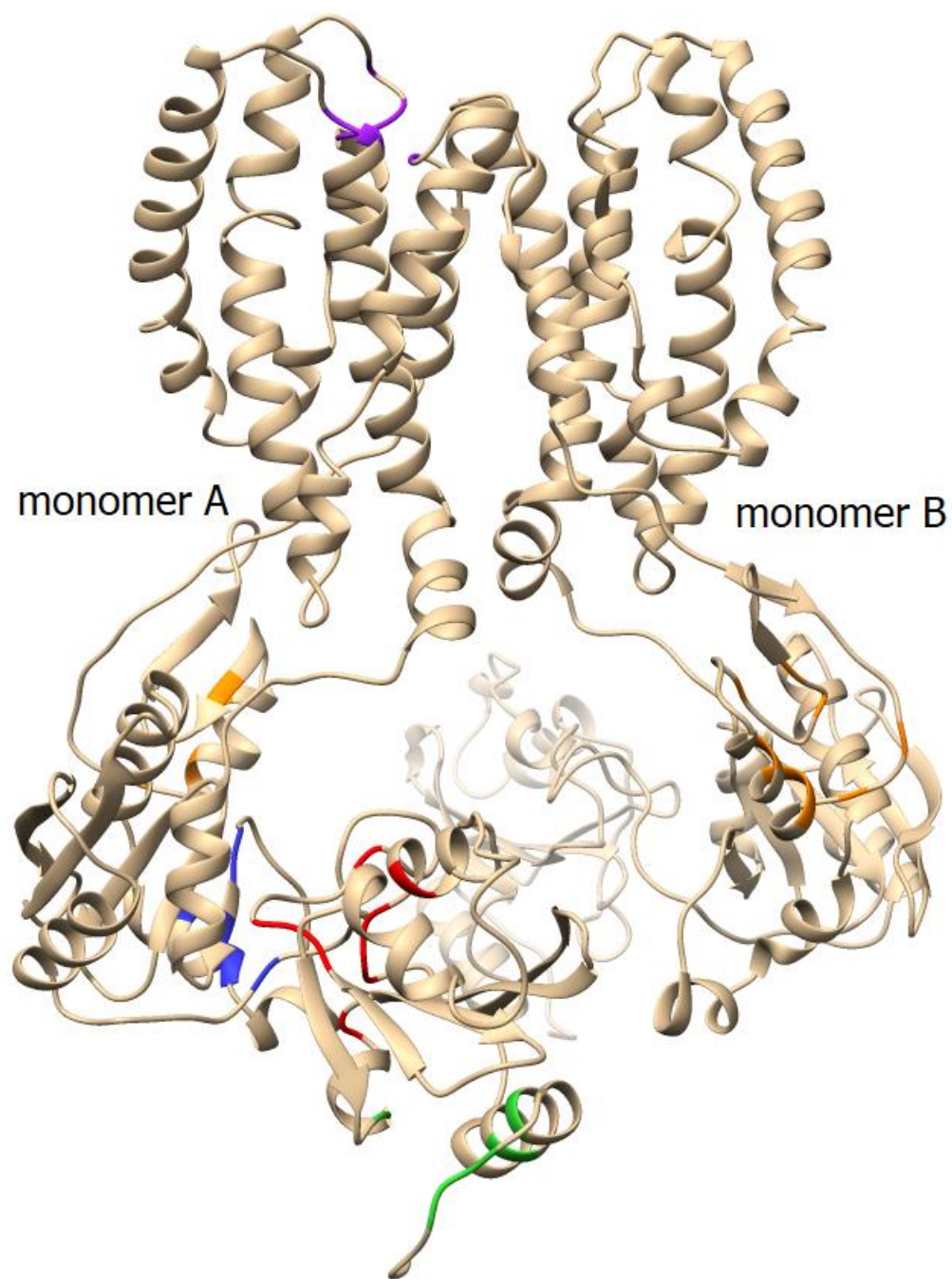
Wyniki dokowania L-alaniny do GerAA		
Numer klastra	Ilość struktur	Reszty aminokwasowe
1	4389	Glu339, Arg433, Thr337, Leu341, Ile338, Val426, Thr429, His430, Phe288
2	9925	Asn452, Ans473, Pro474, Pro472, Asn470, Val469, Leu460, Ser456, Asn471, Ser449
3	1552	Leu282, Ser286, Arg283, Gly398, Val382, Val379, Leu378
4	3793	Leu300, Ala299, Ile421, Leu348, Pro351, Pro349, Leu352, Arg241
5	5669	Thr114, Thr115, Pro84, Val154, Asp50, Met48, Gly98, Asn97
6	556	Gly419, Leu417, Tyr418, Ile414, Leu422, Phe423, Ile420
7	704	Ala367, Thr315, Tyr297, Ala371, Ala368, Asm372
8	3986	Pro256, Ile224, Gln223, Leu254, Gln225, Ser206, Gly207, Pro135, Ile203
9	1197	Arg194, Asn197, Phe221, Ile198, Asp454, Asn471, Thr453
Wyniki dokowania L-waliny do GerAA		
Numer klastra	Ilość struktur	Reszty aminokwasowe
1	5744	Asp50, Thr115, Val54, Leu113, Thr114, Ala82, Pro84, Met48, Lys119, Lys118, Thr49, Asn97
2	14273	Ans473, Ser449, Asn452, Val469, Pro475, Leu460, Phe447, Ser456, Asn197, Ala193, Asm471, Ile203, Pro220, Ile198, Gln199, Thr453
3	1227	Phe442, Glu330, Tyr441, Ser479, Leu333, Ser439
4	1461	Leu300, Ala299, Gly419, Ile421, Leu348, Pro351, Pro349
5	1108	Ser206, Ile203, Ile224, Leu212, Gly207, Pro135, Ser219
6	2233	Leu123, Glu47, Thr124, Thr120, Glu129, Lys30, Tyr118, Thr127
7	2503	Asn13, Leu16, Lys20, Thr91, Phe6, Val17, Thr4, Glu5

6.5.1.3. Wyniki dokowania germinantów do dimeru GerAA

Serwer PLIP sugeruje, że dimer GerAA może mieć aż osiem potencjalnych miejsc wiążących ligand: 1) Leu14A, Ala15A; 2) Arg151A, Ser155A, Leu159A; 3) Leu307A, Leu308A; 4) Asp12B, Leu14B, Ala15B; 5) Arg151B, Arg153B, Ser155B, Leu159B; 6) Leu307B; 7) Ile338A, Val426A, Arg433A, Gln434A; 8) Ile 338B, Leu341B, Arg342B, Val436B, Arg433B, Gln434B. Są to miejsca rozłożone symetrycznie na obu monomerach GerAA oraz odpowiadają one kieszeniom wiążącym przewidzianym dla monomeru tego białka. Podobnie jak w przypadku monomeru, są to miejsca znajdujące się poza wewnętrzną błoną przetrwalnika.

Serwer BiteNet sugeruje trzy potencjalne kieszenie wiążące: 1) Asn13A, Leu16A, Val17A, Lys20A, Lys89A, Thr91A, Ile92A; 2) Glu5B, Phe6B, Lys16B, Val17B, Lys89B, Thr91B, Ile92B, Phe130B; 3) Phe284A, Tyr428A, Arg280B, Phe284B, Tyr428B oraz Thr432B. Trzeci z przewidzianych regionów obejmuje reszty aminokwasowe obu monomerów, sugerując, że wiązanie germinantów może się odbywać na styku monomerów.

Wyniki przeprowadzonego przeze mnie dokowania globalnego zebrane w tabeli 8 pokazują trzy regiony, do których dokowały się oba germinanty i kilka rozbieżnych regionów, aczkolwiek wiele z opisanych w tabeli klastrów nachodzi na siebie. Przykładowo lokalizacja reprezentowana przez klastry 6 oraz 4 L-alaniny występuje tylko dla L-alaniny — klaster 4 (kolor czerwony na rysunku 44) jest licznie reprezentowany, natomiast klaster 6 (kolor fioletowy) obejmuje reszty 299 oraz 300, które wchodzi w skład helisy 273-302. Klaster 2 dla L-alaniny oraz klaster 5 dla L-waliny są najliczniej reprezentowane spośród wszystkich regionów i na rysunku zostały oznaczone kolorem niebieskim. Klastry numer 1 obu germinantów zaznaczone zostały kolorem zielonym. Natomiast klastry 3 i 8 dla L-alaniny wraz z grupami 5 i 6 dla L-waliny są do siebie bardzo podobne, z tego powodu na rysunku 44 zaznaczyłam je wspólnie — kolorem pomarańczowym. Jak widać na rysunku 44, w przypadku dimeru GerAA germinanty najczęściej dokowały się do regionów zlokalizowanych w niebłonowych częściach dimeru.



Rysunek 44: Rozmieszczenie potencjalnych miejsc wiązania germinantów do dimeru GerAA

Tabela 8: Wyniki dokowania globalnego germinantów do dimeru GerAA.

Wyniki dokowania globalnego L-alaniny do dimeru GerAA		
Numer klastra	Ilość struktur	Reszty aminokwasowe
1	1068	Thr91A, Lys7A, Phe6A, Glu5A, Asn13A, Leu16A
2	10429	Ile203A, Asn266A, Pro135A, Ile224A, Gly207A, Gln223A, Ser219A
3	3195	Asn452B, Ser449B, Ser456B, Pro475B, Leu460B, Val469B, Ile457B
4	4676	Met48A, Gly98A, Gly116A, Thr115A, Thr114A, Asn97A, Thr49A, Arg121A, Asp50A, Val54A
5	3215	Glu5B, Thr4B, Lys20B, Val17B, Leu16B
6	1803	Pro351A, Pro349A, Ala299A, Leu300A, Leu348A, Ile421A
7	1286	Gly353B, Tyr418A, Leu348B, Pro349B, Ile421B, Leu358B, Pro351A, Leu352B
8	2420	Asn473B, Val469A, Leu460A, Asn452A, Ser456A
Wyniki dokowania globalnego L-waliny do dimeru GerAA		
Numer klastra	Liczba struktur	Reszty aminokwasowe
1	2362	Thr91A, Lys90A, Ile92A, Asn13A, Lys7A, Phe6A
2	1431	Arg402A, Thr435A, Leu431A, Tyr428A, Thr432A, Phe406B
3	1456	Glu330B, Ser479B, Phe442B, Val384B, Leu333B, Ser349B, Thr441B
4	1513	Ser456A, Asn452B, Ile457B, Pro475B, Ser449B, Val469B, Leu460B
5	3554	Ile224A, Ser206A, Gln223A, Asn226A, Pro135A, Gly207A, Ile203A
6	2888	Asn452A, Asn197A, Ser449A, Asn473A, Pro472A, Phe448A, Ala193A, Leu460A
7	1436	Arg402B, Phe406B, Leu431A, Tyr428A, Thr432A, Thr435A

6.5.2. Dokowania germinantów do podjednostki B

6.5.2.1. Wyniki przewidywań miejsc oddziaływań z ligandem dla GerAB

Serwer MetaPocket, który został użyty przeze mnie w 2016 roku przy wczesnych pracach nad modelem receptora przewiduje trzy potencjalne kieszenie wiążące dla białka GerAB. Aktualnie serwer jest wyłączony z użytku, dlatego dla modeli kompleksu oraz poprawionego modelu GerAA wykorzystałam serwery BiteNet oraz PLIP. Miejsca wiązania liganda, według tego serwera, składają się z następujących reszt aminokwasów: kieszeń pierwsza — Gln283, Val193, Thr106, Ile196, Trp280; kieszeń druga — Tyr71, Leu207, Trp253, His210, Phe272, Leu10, Ile18, Pro254; oraz trzecia kieszeń wiążąca — Ser195, Leu191, Thr192, Val318, Gly52, Glu51, Thr47. Warto zwrócić uwagę na to, że MetaPocket wskazuje resztę 51 jako potencjalnie oddziałującą z ligandem. Jest to reszta, która według mojego modelu receptora GerA zbudowanego na dimerach bierze udział w interakcjach pomiędzy podjednostkami. Z kolei w literaturze opisywana jest jako istotna dla poprawnego funkcjonowania receptora.

Serwer PLIP z kolei sugeruje, że białko GerAB nie posiada miejsc wiążących — prawdopodobnie wynika to z niedostosowania serwera do przewidywań miejsc interakcji z białkami błonowymi. Również w przypadku białka GerAA wskazywał jedynie regiony znajdujące się poza obszarem błonowym.

Serwer BiteNet dla podjednostki B przewiduje jedno miejsce wiązania: Arg33, Ala34, Thr106, Met 109, Val279 oraz Trp280. W wynikach dokowania globalnego reszta Arg33 obserwowana jest przeze mnie jako reszta aminokwasowa biorąca udział w interakcji z germinantami w kompleksie dimerów GerAB, natomiast reszty Thr106 i Trp280 — w interakcjach z monomerem GerAB.

6.5.2.2. Dokowanie globalne do monomeru GerAB

Wyniki globalnego dokowania L-alaniny oraz L-waliny do podjednostki B receptora GerA zebrałam w tabeli 9 oraz na rysunku 45. Na podstawie tych wyników można wytypować siedem regionów, do których dokowała się L-alanina oraz 6 regionów, w których wiązała się L-walina. Część reszt biorących udział w wiązaniach ligandów pokrywa

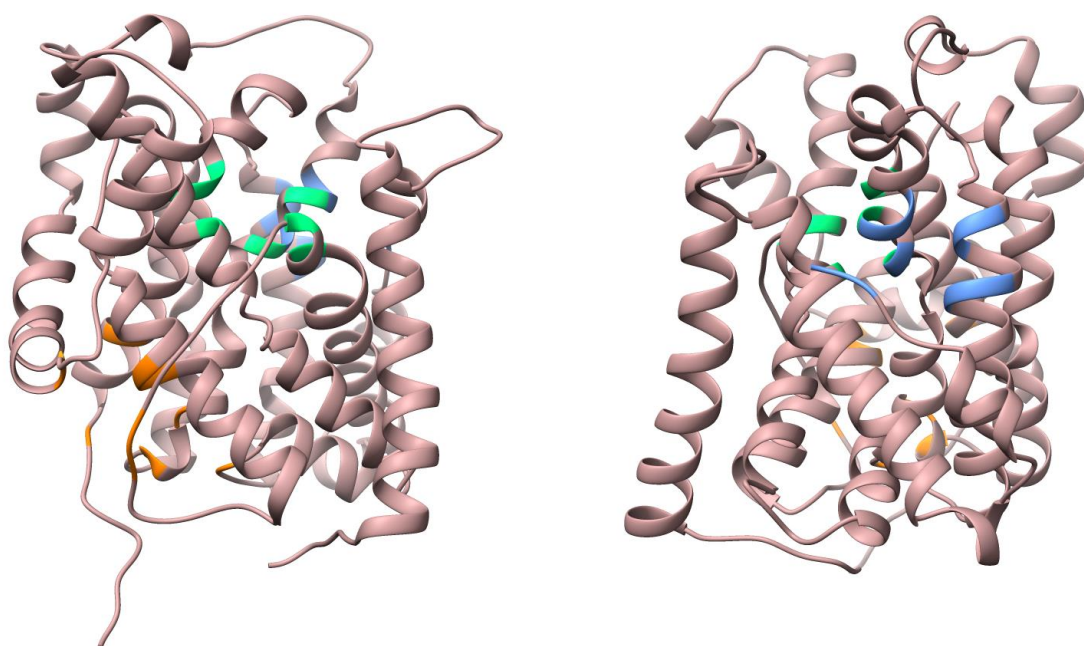
się z resztami aminokwasowymi typowanymi przez serwer MetaPocket jako te, które wchodzą w skład kieszeni wiążących tego białka.

Pierwszy klaster dla L-waliny oraz trzecia kieszeń wiążąca przewidziana przez MetaPocket mają wiele wspólnych reszt aminokwasów: Glu51, Thr47, Ser195, Gly52, Thr192 i Leu191. W tym regionie L-walina dokowała się najczęściej. Z kolei pierwszy klaster L-alaniny praktycznie pokrywa się z pierwszą kieszenią wiążącą wskazaną przez MetaPocket: Thr106, Gln283, Val193, Trp280, Ile196. Region dokowań przewidziany przeze mnie podczas dokowania globalnego jest większy o kilka reszt aminokwasów. Do tego regionu L-alanina dokowała się najwięcej razy pośród wszystkich zaobserwowanych klastrów. Druga kieszeń wiążąca z serwera pokrywa się z klastrem 5 dla L-alaniny oraz 3 dla L-waliny. W sumie do obu klastrów germinanty dokowały się 56 709 razy.

Reszty aminokwasowe Thr106 oraz Trp280 wskazane w klastrze pierwszym opisującym wyniki dokowania L-alaniny do GerAB zostały przewidziane jako część potencjalnej kieszeni wiążącej również przez serwer BiteNet.

Tabela 9: Wyniki dokowania globalnego germinantów L-alaniny i L-waliny do monomeru GerAB.

Numer klastra	Liczba struktur	Reszty aminokwasowe
1	159350	Thr106, Thr30, Leu31, Gln283, Trp280, Val193, Ile196
2	25563	Glu209, Gln5, Gln70, Phe268, Ile269, Glu270
3	15144	Thr224, Ala245, Arg107, Leu29, Ile135, Ala108, Lys249, Val248, Glu111
4	10493	Phe58, Leu61, Asn62, Thr63, Phe57
5	11520	Phe272, Trp253, Leu142, Pro254, Glu273
6	2801	Gln260, Leu99, Gly96, Leu92, Cys95
7	135	Cys133, Tyr136, Met138, Val139
Numer klastra	Liczba struktur	Reszty aminokwasowe
1	145418	Ser195, Leu191, Thr192, Val318, Glu51, Gly52, Phe319, Leu48, Thr47
2	5032	Ala254, Thr244, Arg107, Lys249, Val248, Ile135, Glu111
3	45189	Tyr71, Ile18, Leu10, Pro254, Trp253, Phe272, Leu207, His210
4	7897	Phe58, Leu61, Thr63, Phe57, Asn62, Thr103
5	10152	Gln260, Leu92, Leu347, Phe346, Leu99, Gly96, Cys95
6	11307	Glu209, Gln70, Gln5, Ile269, Lys365, Phe268



Rysunek 45: Trzy kieszenie wiążące wytypowane przez dokowania globalne obu germinantów, MetaPocket oraz BiteNet naniesione na model struktury GerAB. Kieszeń wiążąca zaznaczona kolorem zielonym odpowiada klastrowi pierwszemu z dokowań L-alaniny oraz kieszeni wiążącej numer dwa z MetaPocket.

Kolorem niebieskim oznaczony został klaster numer jeden z dokowań L-waliny oraz pierwszą kieszeń wiążącą MetaPocket. Kolorem pomarańczowym natomiast zaznaczona jest kieszeń wiążąca numer trzy oraz odpowiadające jej klastry 3 (L-walina) i 5 (L-alanina).

6.5.2.3. Dokowanie globalne do dimeru GerAB

Wyniki przewidywań miejsc interakcji uzyskane za pomocą serwera BiteNet wskazują trzy potencjalne kieszenie wiążące dimeru GerAB. Dwie kieszenie (pierwsza i druga) są do siebie symetryczne, natomiast trzecia kieszeń umiejscowiona jest na styku monomerów. Pierwsza przewidziana kieszeń: Arg33A, Thr106A, Met109A, Ala110A i Val279A. Druga kieszeń: Arg33B, Ala34B, Thr106B, Met109B, Val279B, Trp280B. Trzecia, ostatnia, przewidziana kieszeń wiążąca zbudowana jest z następujących reszt aminokwasów: Ala296A, Gly292B, Tyr293B, Ala269B. Dwie pierwsze kieszenie zostały także przewidziane jako miejsca wiązania monomeru GerAB.

Wyniki globalnego dokowania L-alaniny do dimeru GerAB, które przeprowadziłam za pomocą narzędzia AutoDock Vina wskazują dziesięć potencjalnych miejsc, które mogą być odpowiedzialne za interakcje tego białka z germinantami. Wynik dokowania L-waliny z kolei sugeruje dziewięć potencjalnych kieszeni wiążących. Większa część reszt aminokwasowych biorących udział w wiązaniach wodorowych i oddziaływaniach

hydrofobowych regionów z ligandami praktycznie pokrywa się dla obu germinantów, co więcej, pomiędzy kolejnymi klastrami można zauważyć podobieństwa w ich lokalizacji. Wyniki dokowań opisujące reszty aminokwasowe, które brały udział w wiązaniach w danym regionie oraz liczbę kompleksów w poszczególnych klastrach zostały zebrane w tabeli 10.

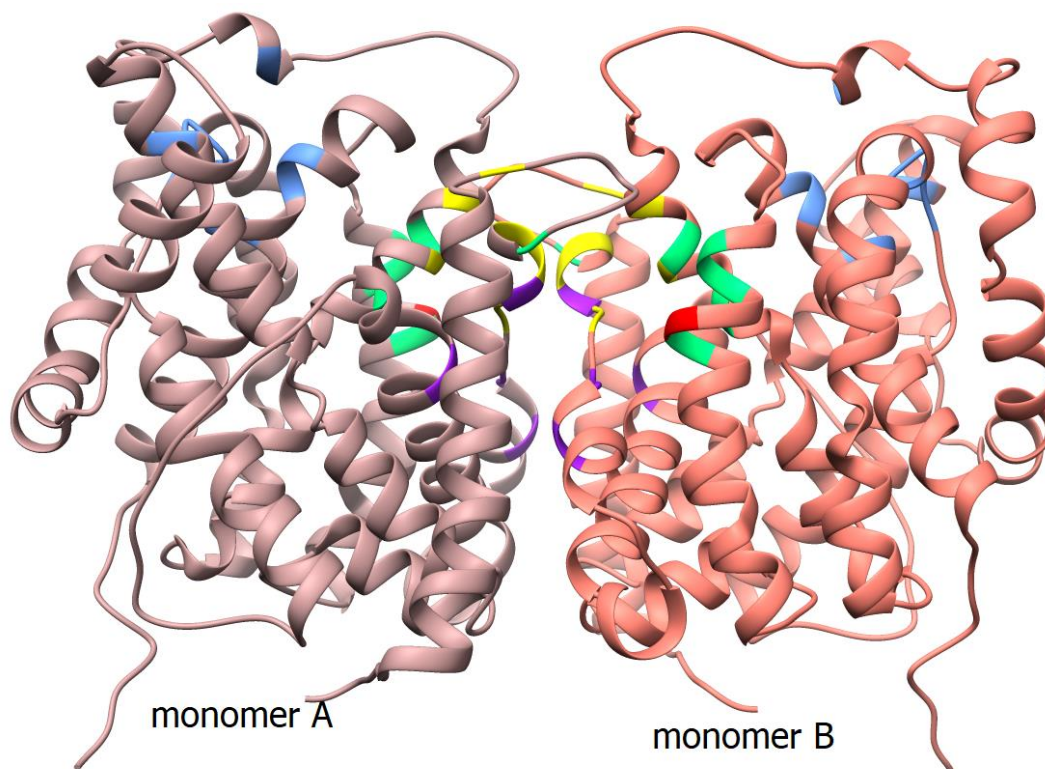
Na podstawie wyników z tabeli 10 można wytypować osiem głównych regionów, które dokowanie globalne wskazuje jako potencjalne kieszenie wiążące. Przy czym, niektóre z tych regionów w dużym stopniu na siebie zachodzą, a część z nich jest identyczna zarówno dla L-alaniny jak i L-waliny. Klastry 1 oraz 4 L-alaniny odpowiadają klastrów 7 i 8 L-waliny i na rysunku 46 zaznaczone są kolorem niebieskim. Klastry 2 oraz 3 L-alaniny odpowiadają klastrów 1 i 2 L-waliny, a na rysunku 46 zostały oznaczone kolorem zielonym — te klastry obejmują miejsca na styku monomerów. Klastry 7 i 8 L-alaniny opisują podobny region jak klastry oznaczone kolorem zielonym. Podobnie w przypadku klastrów 3 i 4 L-waliny, których reszty aminokwasowe, nie należące do klastrów oznaczonych zielonym zostały wyróżnione kolorem czerwonym. Te klastry są bardzo licznie reprezentowane w wynikach dokowań L-waliny. Dodatkowo, kolorem fioletowym oznaczone zostały klastry 5 i 6 L-waliny, a kolorem żółtym klastry 5 i 9 L-alaniny.

Rysunek 46 pokazuje rozłożenie potencjalnych miejsc wiązania na strukturze dimeru. Jak widać, reszty aminokwasowe typowane w kilku różnych klastrach zebranych w tabeli 10 budują cztery główne regiony. Najwięcej kompleksów opisujących wyniki dokowania znajduje się regionach umiejscowionych na styku obu monomerów GerAB, przy czym wydaje się, że L-walina może dokować się w głębszych warstwach dimeru niż L-alanina. Wyniki dokowań do dimeru są bardzo podobne do wyników dokowań dla monomeru GerAB.

Reszta Glu51, wspomniana już wielokrotnie, zarówno w kontekście interakcji GerAB z podjednostkami jak i w interakcji z germinantem, jest resztą, której mutacje skutkują nieaktywnym receptorem GerA (Moir, Lafferty i Smith, 1979). Arg33 obecna w klastrach obu germinantów jest wskazywana przez serwer BiteNet jako jedna z reszt oddziałujących z ligandem.

Tabela 10: Wyniki dokowań globalnych L-alaniny oraz L-waliny do dimeru podjednostek B receptora GerA.

Wyniki dokowania L-alaniny do dimeru GerAB		
Numer klastra	Liczba struktur	Reszty aminokwasowe
1	3838	Asn174A, Gly241A, Leu29A, Leu28A, Leu243A, Lys114A, Pro32A, Arg33A, Val240A, Val239A, Ala242A
2	10442	Ser195A, Thr192A, Leu48A, Gly52A, Leu191A, Trp324B, Thr47A
3	8033	Ser195B, Thr192B, Leu48B, Gly52B, Leu191B, Trp324A, Thr47B
4	5784	Asn174B, Gly241B, Leu29B, Leu28B, Leu243B, Lys114B, Pro32B, Arg33B, Val240B, Val239B, Ala242B
5	1334	Asn189B, Ser322B, Thr192B, Val318B, Phe321A, Phe331B, Phe319B, Phe319A
6	1481	Ser195A, Val318A, Ile316A, Thr317A, Trp324B, Phe288A, Thr192A, Phe319A
7	1261	Arg33A, Leu28A, Leu29A, Thr30A, Pro32A, Lys114A, Val113A
8	609	Arg33B, Leu28B, Leu29B, Thr30B, Pro32B, Lys114B, Val113B
9	641	Asn189A, Ser322A, Thr192A, Val318A, Phe321B, Phe331A, Phe319A, Phe319B
10	502	Glu209B, Gln5B, Ile269B, Gln270B, Phe268B
Wyniki dokowania globalnego L-waliny do dimeru GerAB		
Numer klastra	Liczba struktur	Reszty aminokwasowe
1	5062	Thr47A, Ser195A, Thr192A, Leu191A, Gly52A, Leu48A, Glu51A, Trp324B, Ile318A, Ile56A
2	6717	Thr47B, Ser195B, Thr192B, Leu191B, Gly52B, Leu48B, Glu51B, Trp324A, Ile318B, Ile56B, Phe288B
3	8480	Ile316A, Gly52A, Glu51A, Trp324B, Tyr377B
4	9504	Ile316B, Gly52B, Glu51B, Trp324A, Tyr377A
5	1341	Tyr320A, Ile313B, Ile316B, Phe285A, Thr319B, Gly315B
6	1292	Tyr320B, Ile313A, Ile316A, Phe285B, Thr319A, Gly315A
7	2286	Asn174A, Gly241A, Leu29A, Thr30A, Val240A, Lys114A, Val113A, Arg33A, Pro32A, Val239A
8	2684	Asn174B, Gly241B, Leu29B, Thr30B, Val240B, Lys114B, Val113B, Arg33B, Pro32B, Val239B
9	523	Gly52B, Leu191B, Trp324A, Glu51B, Leu48B, Thr47B



Rysunek 46: Rozmieszczenie potencjalnych miejsc wiążących dimeru GerAB

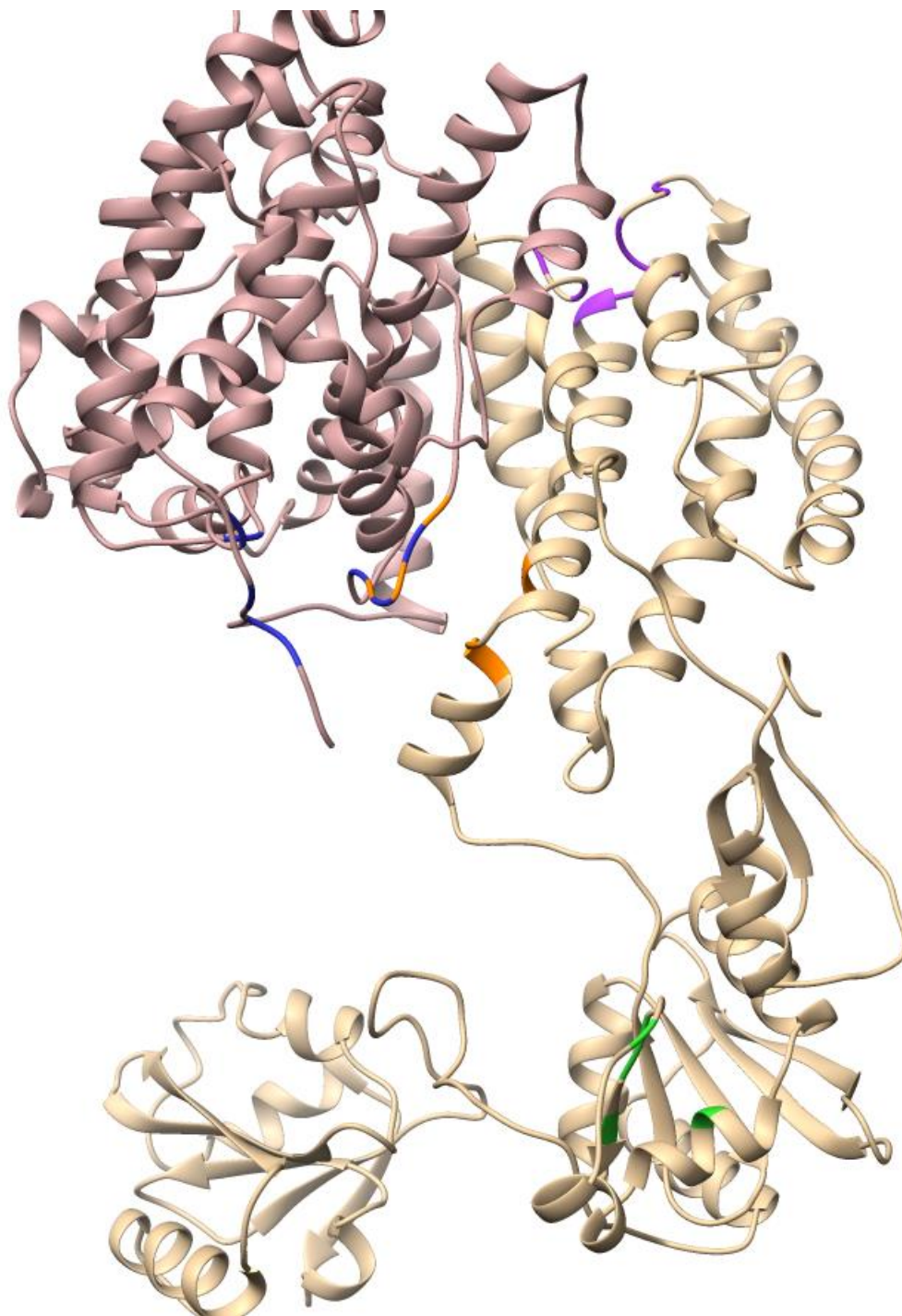
6.5.3. Oddziaływania germinantów z kompleksem receptora GerA zbudowanym z monomerów

Serwer BiteNet przewiduje trzy potencjalne kieszenie wiążące dla kompleksu zbudowanego z monomerów GerAA i GerAB. Pierwsza kieszeń jest skupiona na białku GerAA: Asn13A, Leu16A, Val17A, Lys20A, Lys89A, Thr91A, Ile92A. Druga kieszeń wskazuje trzy reszty aminokwasowe białka GerAB: Met109B, Leu128B, Trp280B, natomiast trzeci przewidziany region znajduje się w białku GerAA: Ans24A, Asp26A, Leu27A, Phe45A, Ser46A, Ile95A, Leu96A.

Porównując wyniki przewidywań serwera BiteNet z wynikami dokowań globalnych dla kompleksu GerA zbudowanego z monomerów, nie widać podobieństw pomiędzy wytypowanymi przez te metody resztami aminokwasowymi. Natomiast przewidziane przez BiteNet pozycje przewijają się podczas wyników dokowań do dimerów i monomerów białek GerAA i GerAB. Należy zwrócić uwagę na to, że kieszenie 1 i 3 w całości znajdują się w niebłonowej domenie białka GerAA, podczas gdy uzyskane przeze mnie wyniki dokowań

globalnych wskazują również regiony błonowe tego białka, jako potencjalne miejsca wiązania germinantów. Trzecia kieszeń wiążąca przewidziana przez BiteNet zawiera z kolei trzy reszty aminokwasowe, które odpowiadają miejscu wiązania L-waliny i glukozy wskazanym w homologue białka GerAA: Asp26, Phe45 oraz Leu96.

Dokowania globalne germinantów wskazują osiem miejsc (wyniki opisane zostały w tabeli 11), do których L-alanina dokowała się najczęściej oraz cztery takie miejsca dla L-waliny. Porównanie klastrów obu germinantów pokazuje, że wszystkie regiony, do których dokowała się L-walina występują także w przypadku dokowań L-alaniny. Klaster 1 L-waliny odpowiada klastrowi 7 L-alaniny i zlokalizowany jest w całości w białku GerAB (rysunek 47, kolor niebieski). Do tych klastrów germinanty dokowały się najczęściej. Drugimi pod względem częstotliwości dokowań są klastry 3 dla L-alaniny oraz 2 dla L-waliny. Ten region zlokalizowany jest na styku obu monomerów, a na rysunku 47 został zaznaczony kolorem zielonym. Trzeci klaster dla L-waliny i 2 dla L-alaniny obejmuje reszty aminokwasowe z białka GerAA i na rysunku 47 zaznaczony jest kolorem pomarańczowym. Ostatni wspólny region opisywany jest przez klastry numer 4 obu germinantów (kolor fioletowy). Jest to region zahaczający o reszty 299 i 300 białka GerAA będące fragmentem helisy 273-302 opisanej w podrozdziale 6.3.1 *Pozycje 299 i 302*.



Rysunek 47: Miejsca wiązania wspólne dla obu germinantów (L-alaniny oraz L-waliny) wytypowane podczas dokowania globalnego.

Tabela 11: Wyniki dokowania germinantów do kompleksu GerA zbudowanego z monomerów GerAA i GerAB.

Wyniki dokowania L-alaniny do kompleksu monomerów GerA		
Numer klastra	Liczba struktur	Reszty aminokwasowe
1	507	Leu282A, Ser288A, Leu378A, Gly298A, Arg283A, Val379A, Val382A
2	3691	Arg280A, Ile269B, Arg271A, Phe272B, Glu273B, Ser274B, Arg271B
3	3603	Ile203A, Asn226A, Gly207A, Pro135A, Ile224A
4	2028	Pro351A, Pro349A, Leu348A, Ile241A, Leu300A, Ala299A
5	1108	Asn452A, Leu460A, Asn473A, Val469A, Ser456A
6	558	Val426A, Thr429A, Ile338A, Glu339A, Leu341A, Phe388A
7	4308	Glu5B, Gln209B, Tyr71B, Phe268B, Pro7B, Lys4B, His210B, Gln5B, Lys4B
8	1189	Thr115A, Thr114A, Arg121A, Val54A, Asp50
Wyniki dokowania L-waliny do kompleksu monomerów GerA		
Numer klastra	Liczba struktur	Reszty aminokwasowe
1	5322	Gln5B, Lys4B, His210B, Glu270B, Phe272B, Phe268B, Tyr71B, Pro7B
2	534	Ile203A, Glu223A, Ser219A, Ile224A, Asn226A
3	5439	Arg280A, Glu273B, Arg271B, Ile269B, Phe272B, Ile247A, Ser275A
4	1151	Leu352A, Leu348A, Pro351A, Pro349A, Gly419A, Leu300A, Ile421A, Ala299A

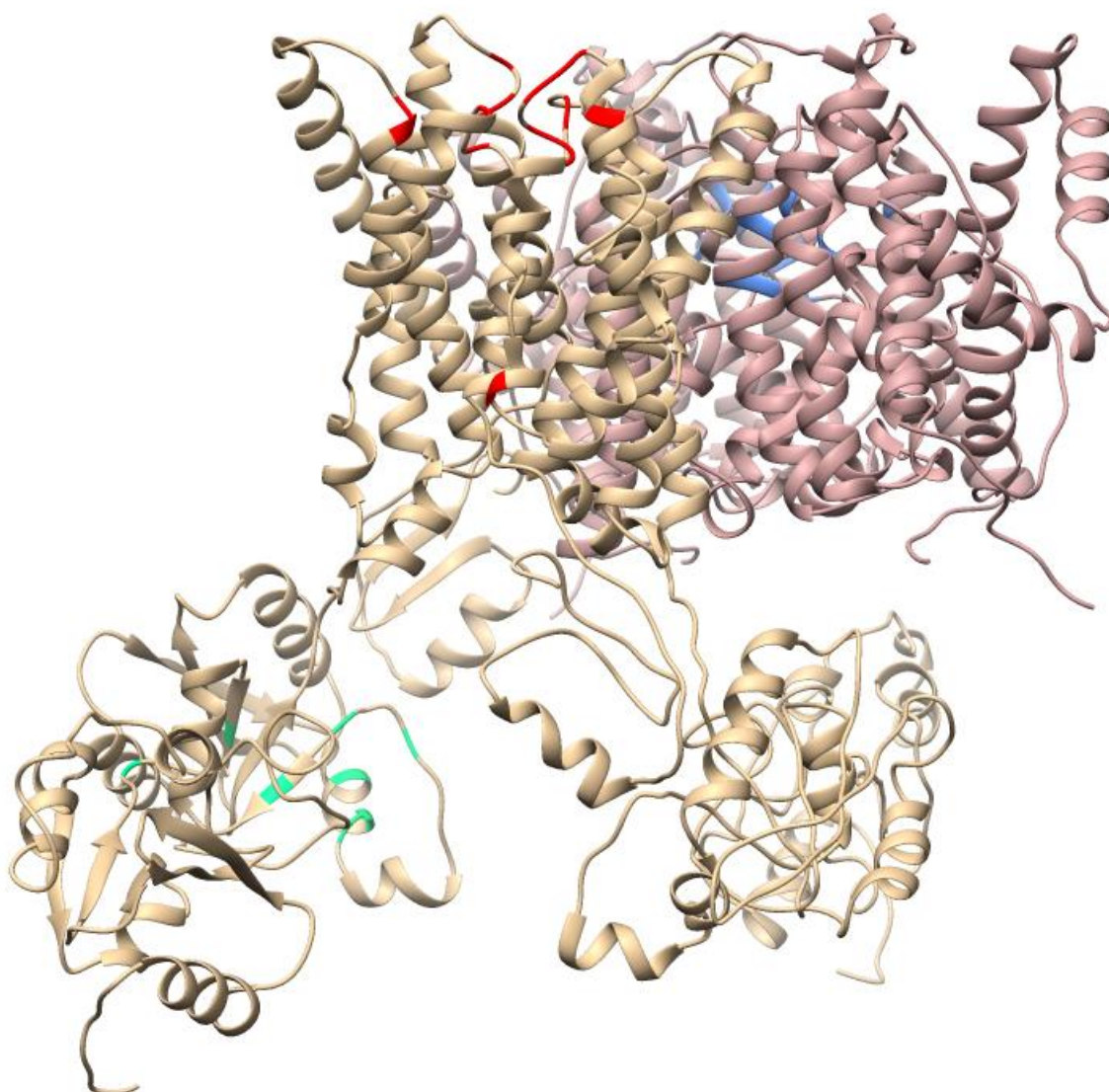
6.5.4. Oddziaływania germinantów z kompleksem receptora GerA zbudowanym z dimerów

Serwer BiteNet przewiduje osiem potencjalnych kieszeni wiążących dla kompleksu zbudowanego z dimerów GerAA i GerAB. Łańcuchy A oraz B to monomery podjednostki B, natomiast łańcuchy C i D reprezentują monomery białka GerAA. Miejsca przewidziane przez BiteNet znajdują się zarówno w części błonowej kompleksu jak i części niebłonowej — dokładnie w dwóch pozabłonowych domenach GerAA. Kieszenie wiążące przewidziane przez BiteNet są następujące:

1. Asn13C, Leu16C, Val17C, Lys20C, Lys89C, Thr91C, Ile92C
2. Leu352A, Phe353A, Thr307B, Lys308B, Thr309B, Tyr270D, Glu271D, Ile274D, Met397D
3. Ile341A, Glu51B, Phe55B, Ile315B, Ile316B, Arg405D
4. Ser310B, Thr435C, Arg283D, Arg402D
5. Glu5D, Phe6D, Lys89D, Lys90D, Thr91D, Ile92D, Phe103D
6. Ile299B, Ile230B, Leu293D, Ser294D
7. Phe350A, Leu333C, Met334C, Gly438C, Ser439C, Tyr441C
8. Arg33B, Ala34B, Thr106B, Met109B, Ala110B, Val113B, Trp280B

Już na tym etapie łatwo zauważyć, że część kieszeni wiążących powtarza się od przewidywań monomerów obu podjednostek, przez przewidywania dla dimerów, aż do przewidywań kieszeni wiążących dla kompleksu. Regiony te warte są wzięcia pod uwagę w rozważaniach dotyczących umiejscowienia kieszeni wiążącej i interakcji z germinantami w badaniach *in vivo*. Ta powtarzalność wyników może również wynikać z tego, że są to miejsca dość łatwo dostępne dla potencjalnego liganda.

Tabela 12 zbiera wyniki dokowań globalnych, które przeprowadziłam przy pomocy AutoDock Vina dla obu germinantów i kompleksu dimerów. Klastry 2 i 3 L-alaniny odpowiadają klastrom 1 i 2 L-waliny (kolor niebieski na rysunku 48), klastry numer 5 obu germinantów obejmują ten sam region (kolor zielony), natomiast klastry 8 i 10 dla L-alaniny odpowiadają klastrowi 6 L-waliny (kolor czerwony). Dodatkowo ligandy te dokują się też do klastrow, które nie znajdują swoich odpowiedników. Taki rozkład kompleksów może sugerować, że oba germinanty będą dokować się do któregoś ze wspólnych regionów (są to regiony przewijające się przez dokowania do monomerów i dimerów), nie należy jednak odrzucać możliwości, według której L-alanina i L-walina dokować się mogą do innych miejsc wiążących.



Rysunek 48: Umieszczenie najczęściej występujących regionów dokowania L-alaniny i L-waliny do kompleksu dimerów GerA. Najczęściej germinanty dokowały się do miejsc oznaczonych kolorem niebieskim i znajduje się on w monomerach podjednostki B. Regiony zielony i czerwony znajdują się w strukturze białka GerAA.

Tabela 12: Wyniki dokowań globalnych L-alaniny oraz L-waliny do kompleksu GerA zbudowanego z dimerów GerAA i GerAB.

Wyniki dokowania L-alaniny do kompleksu dimerów GerA		
Numer klastra	Ilość struktur	Reszty aminokwasowe
1	1144	Thr30A, Leu29A, Arg33A, Pro32A, Val240A, Asn174A
2	547	Thr30B, Leu29B, Arg33B, Pro32B, Val240B, Asn174B
3	1723	Gly52B, Ser195B, Thr47B, Leu48B, Leu191B, Thr192B, Trp324A
4	2068	Gly52A, Ser195A, Thr47A, Leu48A, Leu191A, Thr192A, Trp324B
5	1255	Ile203C, Asn226C, Pro135C, Ile224C, Gly207C
6	2213	Thr480C, Asp478C, Gly481C, Leu264A
7	2098	Thr302B, Thr307, Ser306B, Pro266D, Tyr270D, Asp267D, Thr309B, Leu352A, Phe303B
8	508	Pro351C, Pro349C, Leu421C, Leu300C, Ala299C
9	869	Ser322A, Phe319A, Tyr340A, Val193A, Thr192A, Leu323A, Phe288A, Phe284A, Gln283A
10	1392	Gly335D, Tyr418C, Pro349C, Pro349D, Ile421D, Leu348D, Pro351D, Leu358D, Leu300C
11	434	Glu370C, Gln366C, Val336C, Phe325C, Glu339C, Leu373C
Wyniki dokowania L-alaniny do kompleksu dimerów GerA		
Numer klastra	Ilość struktur	Reszty aminokwasowe
1	1356	Ser195B, Thr47B, Leu48B, Leu191B, Glu51B, Gly52B, Trp324A
2	2097	Ser195A, Thr47A, Leu48A, Leu191A, Glu51A, Gly52A, Trp324B
3	1119	Ile316A, Glu51A, Gly52A, Trp324B, Tyr337B
4	808	Glu370C, Gly365C, Gln366C, Glu339C, Val135A, Val336C, Phe325C, Leu373C
5	1790	Ile244C, Ile203C, Asn226C, Ser206C, Ser219C, Gln223C
6	1207	Pro351D, Pro349D, Tyr418C, Leu352D, Gly353D, Gly419D, Leu348D, Leu421D, Leu358D
7	595	Tyr317A, Tyr320B, Ile313A, Gly315A, Ile316A, Phe285B
8	938	Tyr317B, Tyr320A, Ile313B, Gly315B, Ile316B, Phe285A

Jeną z reszt aminokwasowych często występujących w regionie wiązania germinantów jest reszta E335 białka GerAA. Ta reszta jest opisywana w literaturze jako wysoce konserwowana oraz krytyczna dla zachowania poprawnej aktywności receptora GerA (Mongkolthanasakul, Cooper, Mawer, Allan i Moir, 2011). Wyniki dokowania wskazują również region helisy 273-302 oraz reszty Ala299 i Leu300 jako biorące udział w oddziaływaniach z L-alaniną, jak wiemy z poprzednich rozdziałów rozprawy miejsce 299 jest istotne dla odpowiedzi receptora na L-alaninę (Grela i inni, 2018).

W przypadku obu dokowań większość przewidywanych kieszeni wiążących znajduje się po stronie białek B, co może potwierdzać hipotezę postawioną przez (Artzi i inni, 2021) mówiącą o tym, że to GerAB jest odpowiedzialne za interakcje z germinantami. Niemniej, germinanty wciąż wielokrotnie dokowały się również do miejsc znajdujących się na styku obu dimerów oraz do białek GerAA. Rozmieszczenie potencjalnych kieszeni wiążących może sugerować tymczasowy charakter wiązania ligandów — wędrują one w głąb przetrwalnika, perforując błonę i w trakcie swojej wędrówki, oddziałują z obiema podjednostkami, wiążąc się do różnych kieszeni wiążących.

6.5.5. Podsumowanie wyników dokowań globalnych

Uzyskane przeze mnie wyniki dokowań globalnych germinantów do białek receptora GerA pokazują, że kilka regionów potencjalnie wiążących germinanty przewija się przez wszystkie przeprowadzone eksperymenty. Dla podjednostki GerAA są to cztery regiony, natomiast dla podjednostki GerAB można wyróżnić trzy takie regiony — zostały one zebrane w tabeli 13. Fakt, że powtarzają się one we wszystkich dokowaniach, pozwala sądzić, że są najbardziej prawdopodobnymi kieszeniami wiążącymi germinanty.

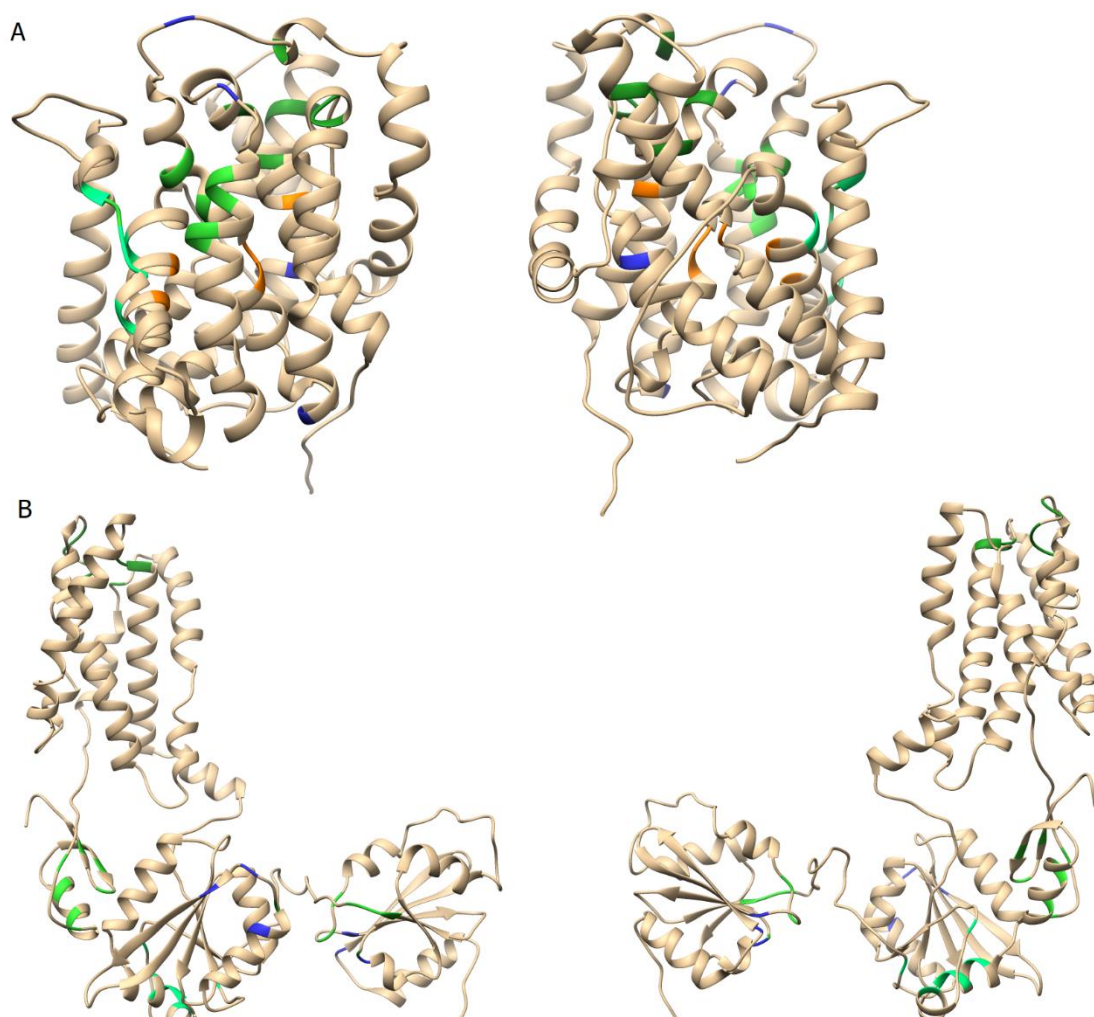
Region GerAA1 zawiera w sobie dwa interesujące miejsca. Po pierwsze jest to reszta 299, w której mutacje mają wpływ na funkcjonowanie receptora oraz reszta 300, która razem z 299 budują helisę błonową białka GerAA (helisa 273-302). Tuż obok nich znajduje się reszta 302, która, co prawda nie pojawiała się w opisie lokalizacji klastrów, ale podobnie jak reszta 299, jest niezwykle istotna dla poprawnego funkcjonowania receptora GerA (Grela i inni, 2018). Drugim aspektem jest typowanie reszt aminokwasowych 348 i 349, jako biorących udział w interakcjach z germinantami. Reszty te znajdują się w bezpośrednim sąsiedztwie reszty Arg347 opisanej w podrozdziale 6.3.2. Mutacje polegające na podstawieniu w miejscu 347 histydyny lub alaniny skutkowały brakiem kiełkowania spory

w obecności L-alaniny. Prawdopodobnie więc zasadowość tego regionu jest istotna dla prawidłowego wiązania L-alaniny do receptora. Co ciekawe, region ten znacznie częściej pojawiał się w dokowaniach L-alaniny niż w dokowaniach L-waliny.

Tabela 13: Tabela zbierająca regiony powtarzające się w wynikach dokowania monomerów, dimerów oraz kompleksów.

Region	Reszty aminokwasowe
GerAA1	Leu300, Ala299, Ile421, Leu348, Pro351, Pro349, Leu352
GerAA2	Asp50, Thr115, Thr114, Val154, Leu113, Met48, Asn97
GerAA3	Ser206, Ile203, Leu212, Gly207, Pro135, Ser219
GerAA4	Asn452, Asn472, Ser469, Pro475, Leu460, Phe447, Ser456
GerAB1	Ser195, Leu191, Thr192, Val318, Glu51, Gly52, Phe319, Leu48, Thr47
GerAB2	Thr30, Arg33, Pro32, Leu29, Val240, Asn174, Gly241, Val113, Lys114, Val239
GerAB3	Tyr317, Tyr320, Ile313, Gly315, Ile316, Phe285

Najnowsze publikacje związane ze strukturą homologa GerAA oraz modelami GerAB wskazują potencjalne kieszenie wiążące pojedynczych podjednostek receptora GerA. Badania przeprowadzone przez Artzi i inni (Artzi i inni, 2021) sugerują, że kieszeń wiążąca germinanty znajduje się w białku GerAB i przypomina ona miejsce wiązania L-alaniny w transporterze GkApcT. Prezentowana w publikacji kieszeń wiążąca składa się z następujących reszt aminokwasowych: G25, V101, Y291, G200, L199, T287. Blinker i inni (Blinker, Vreede, Setlow i Brul, 2021) z kolei wskazuje (na podstawie trzech modeli i trzech potencjalnych kieszeni wiążących) reszty N40, L180, K215 oraz T250 jako miejsca potencjalnie wiążące L-alaninę. Reszty aminokwasowe 195 oraz 258 opisane w tabeli 13, umiejscowione są w bliskim sąsiedztwie kieszeni wiążącej wskazywanej przez badania Artzi i inni.



Rysunek 49: Porównanie najczęściej występujących lokalizacji dokowania germinantów podczas dokowań globalnych z danymi dostępnymi w literaturze. A) Potencjalne kieszenie wiążące dla GerAB. Odcieniami zielonego zaznaczone są reszty przewidywane przez dokowania globalne opisane w tej pracy, kolorem pomarańczowym oznaczono kieszeń wiążącą przewidzianą przez (Artzi i inni, 2021), a kolorem niebieskim reszty przewidziane przez (Blinker, Vreede, Setlow i Brul, 2021) jako biorące udział w wiązaniu L-alaniny. B) Potencjalne kieszenie wiążące germinanty w białku GerAA. Odcieniami koloru zielonego zaznaczono wyniki dokowań opisane w tabeli 13, a kolorem niebieskim reszty wytypowane na podstawie kieszeni wiążącej homologa (Li i inni, 2019).

6.6. Orientacja białek receptora GerA w błonie

Ponieważ receptor GerA jest umiejscowiony w wewnętrznej błonie przetrwalnika, podjąłem próbę ustalenia jakie orientacje w modelu błony przyjmować mogą struktury poszczególnych podjednostek oraz cały receptor. Model błony został przeze mnie przygotowany za pomocą narzędzia Charmm-GUI oraz informacji uzyskanych z badań prof. Banasia oraz dostępnego modelu błon biologicznych *B. subtilis* opisanego w rozdziale 3.9.2

Model błon biologicznych B. subtilis. Orientacje białek GerAA i GerAB oraz kompleksów w błonie obliczyłam za pomocą narzędzi QMbrane oraz serwera PPM 2.0 uznawanego za złoty standard w dziedzinie badań nad białkami błonowymi.

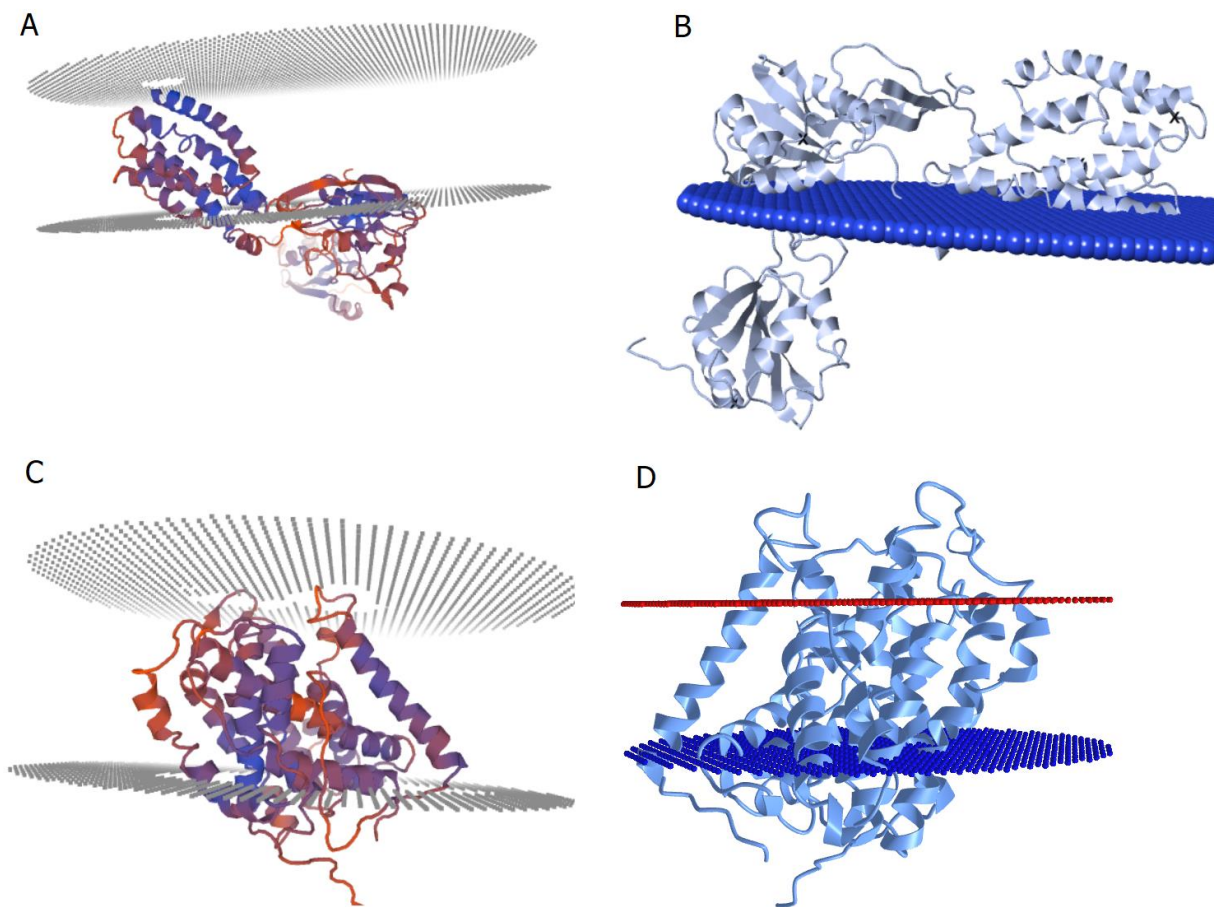
6.6.1. Orientacja monomerów w błonie

Orientacje struktur monomerów białek GerAA oraz GerAB w błonie fosfolipidowej obliczone na podstawie narzędzia QMbrane oraz serwera PPM przedstawione zostały na rysunku 49.

Pseudoenergia potrzebna do zanurzenia białka GerAA w błonie, obliczona za pomocą QMbrane, wynosi $-45296,72$ kcal/mol, co według algorytmu klasyfikuje GerAA jako białko błonowe. Szerokość błony dla tego białka została obliczona na $35,99\text{\AA}$. Serwer PPM natomiast sugeruje, że energia niezbędna dla zanurzenia tego białka w błonie wynosi $-20,4$ kcal/mol, a szerokość błony to w tym przypadku $15,4\pm 4,3\text{\AA}$. Oba narzędzia wskazują, że białko GerAA umiejscowione jest w błonie pod pewnym kątem, PPM sugeruje, że kąt ten wynosi $63^\circ\pm 1^\circ$. Patrząc na panel B na rysunku 49, można zauważyć, że serwer PPM, który uznawany jest za złoty standard w badaniach orientacji białek błonowych, nie poradził sobie zbyt dobrze z umiejscowieniem białka GerAA w błonie. Może to wynikać z faktu, że jest to model teoretyczny, a nie rzeczywista struktura oraz z doboru szablonu podczas modelowania tego białka. Być może korekta części błonowej, o ile w przyszłości dostępne będą nowe informacje dotyczące struktury tej domeny, będzie pomocna w poprawieniu orientacji tej podjednostki w błonie.

Białko GerAB (panele C i D na rysunku 49) według serwera PPM zanurza się w błonie pod kątem $39^\circ\pm 3^\circ$, a szerokość błony wynosi $19,8\pm 0,9\text{\AA}$. Energia niezbędna do zanurzenia podjednostki B w błonie jest równa $-29,9$ kcal/mol. QMbrane wskazuje, że błona wokół białka GerAB może mieć szerokość $33,99\text{\AA}$. Obie wartości wskazują, że jest niewiele szersza od błony otaczającej GerAA. Pseudoenergia potrzebna do przeniesienia białka GerAB w błonę wynosi $-82574,01$ kcal/mol. W przypadku orientacji obliczonej przez PPM większy obszar białka GerAB wystaje poza granice błony, ale oba narzędzia pokazują, że końce tego białka znajdują się wewnątrz rdzenia przetrwalnika.

Różne szerokości błon obliczone dla obu białek mogą wynikać z różnic w rozmiarach i rozmieszczeniu obszarów hydrofobowych białek GerAB oraz GerAA, a także z tego pod jakim kątem algorytmy serwera GerAA umiejscowiły oba białka (rysunek 50 panele B i D).

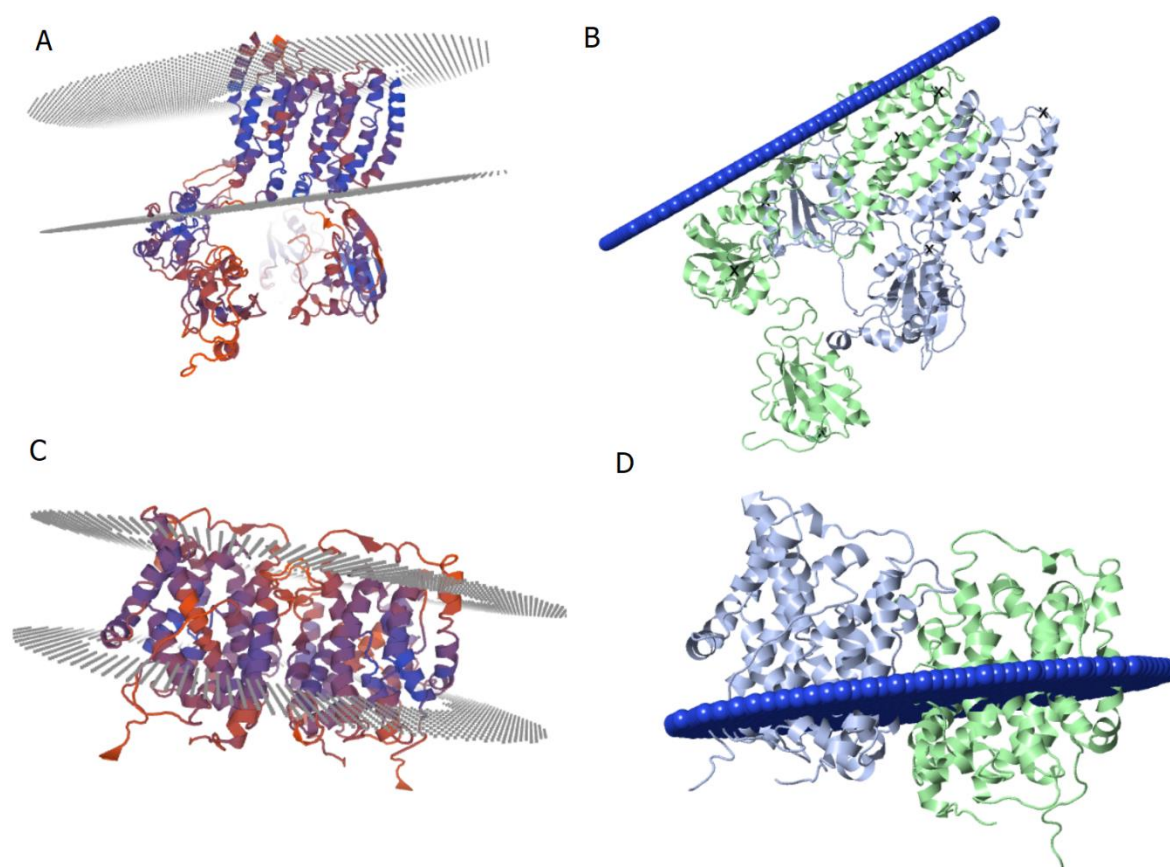


Rysunek 50: Orientacje monomerów GerAA i GerAB w błonie. A) Wyniki obliczeń QMbrane dla białka GerAA. Widzimy, że białko jest zanurzone w błonie pod kątem, a fragment domeny niebłonowej białka znajduje się w obszarze błony. B) Umieszczenie GerAA w błonie według serwera PPM. Ilustracja pokazuje jedynie cytoplazmatyczną granicę błony (kolor niebieski). C) Orientacja GerAB w błonie obliczona na podstawie narzędzia QMbrane. Białko, podobnie jak w sekcji D, zorientowane jest pod kątem, a oba jego końce znajdują się w rdzeniu przetrwalnika. D) Wyniki serwera PPM dla GerAB. Kolorem czerwonym zaznaczona jest zewnętrzna granica błony, a kolorem niebieskim – tzw. cytoplazmatyczna granica. Białko jest zorientowane w błonie pod widocznym kątem. Według obliczeń kąt ten wynosi 39° .

6.6.2. Orientacja dimerów w błonie

QMbrane pokazuje, że model dimeru GerAA otoczony jest błoną o szerokości 36\AA , a pseudoenergia potrzebna na zanurzenie tego dimeru w błonie wynosi $-77891,0$ kcal/mol. Serwer PPM z kolei, podobnie jak w przypadku monomeru GerAA, nie był w stanie umiejscowić dimeru wewnątrz błony, sugerując, że praktycznie całe białko znajduje się w rdzeniu przetrwalnika, a część zanurzona w błonie ma jedynie $6,8\text{\AA}$ szerokości. Szacowana przez PPM energia zanurzenia dimeru w błonie wynosi $-15,2$ kcal/mol.

Dimer GerAB według QMbrane potrzebuje błony o szerokości 24,5Å i w porównaniu do monomeru GerAB, zorientowany jest w błonie pod mniejszym kątem (rysunek 50). Pseudoenergia potrzebna do zanurzenia tego białka wynosi -90107,02 kcal/mol. Serwer PPM, niestety, nie wskazuje przebiegu zewnętrznej granicy błony dla tego dimeru, sugeruje jednak, że szerokość błony wynosi $15\pm 0,7\text{\AA}$, dimer umiejscowiony jest pod kątem $73\pm 2^\circ$, a energia potrzebna na umieszczenie dimeru w błonie wynosi -25,0 kcal/mol.



Rysunek 51: Orientacje dimerów GerAA i GerAB w błonie. A) oraz B) reprezentują obliczone orientacje dla dimeru GerAA. A) Opisuje wyniki obliczeń narzędzia QMbrane, natomiast B) wyniki serwera PPM. W przypadku PPM (zarówno dla GerAA jak i GerAB) widzimy jedynie tzw. cytoplazmatyczną granicę błony. Panel C) pokazuje wyniki obliczenia orientacji dimeru GerAB w błonie za pomocą narzędzia QMbrane, natomiast panel D – wyniki serwera PPM dla tego samego dimeru.

6.6.3. Orientacja kompleksów w błonie

Kompleks błonowej części receptora GerA zbudowany z monomerów podjednostek A i B według narzędzia QMbrane zanurza się w błonie z pseudoenergią równą -115385,79 kcal/mol, a model błony ma szerokość 38Å. Serwer PPM obliczył, że kompleks jest

zorientowany w błonie pod kątem równym $8 \pm 1^\circ$, błona ma szerokość $27,8 \pm 0,3 \text{ \AA}$, a energia transferu kompleksu do błony wynosi $-15,1 \text{ kcal/mol}$.

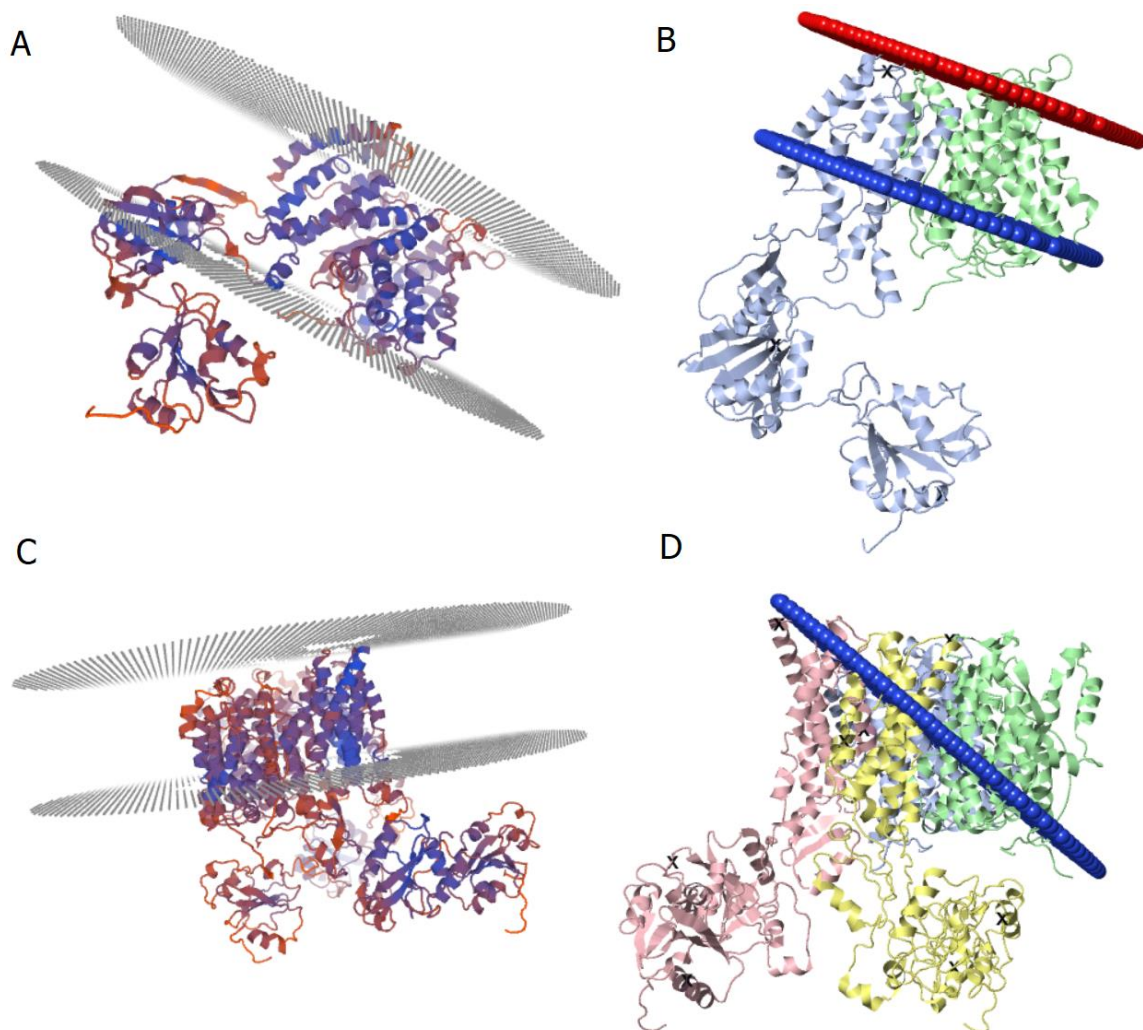
Kompleks dimerów został zorientowany w błonie za pomocą narzędzia QMbrane, serwer PPM z kolei nie poradził sobie z obliczeniem orientacji kompleksu, wskazuje natomiast, że energia potrzebna do zanurzenia tego kompleksu w błonie wynosi $-30,6 \text{ kcal/mol}$. QMbrane wskazuje, że błona wokół transbłonowych części kompleksu ma szerokość 38 \AA , a pseudoenergia niezbędna do umiejscowienia kompleksu w błonie wynosi $-106971,28 \text{ kcal/mol}$.

Porównując ze sobą orientacje obu kompleksów w błonie obliczone przez serwer PPM (rysunek 51), który do tej pory stanowił złoty standard badań nad białkami błonowymi oraz sposób, w jaki kompleksy oddziałują z podjednostką C (rysunek 38 oraz rysunek 35), wydaje się, że model receptora opierający się na monomerach podjednostek jest modelem o wyższej jakości.

Serwer PPM wskazuje 10 regionów transbłonowych dla monomeru GerAB i 6 dla monomeru GerAA w kompleksie GerA zbudowanym z monomerów. Wyniki orientacji białek w błonie pozwoliły mi na zanurzenie kompleksu dimerów GerA oraz monomerów GerA w błonie za pomocą narzędzia Charmm-GUI. Na podstawie danych o orientacji białek w błonie mogłam lepiej opisać orientację receptora GerA podczas budowania modeli kompleks-błona.

Tabela 14: Wyniki przewidywań regionów transbłonowych oraz reszt zanurzonych w błonie na podstawie obliczeń serwera PPM dla kompleksu GerA zbudowanego na monomerach.

Przewidziane przez serwer PPM regiony transbłonowe	
GerAA	1(287- 299), 2(307- 325), 3(338- 351), 4(357- 374), 5(403- 418), 6(419- 429)
GerAB	1(15- 36), 2(44- 64), 3(105- 119), 4(126- 137), 5(148- 171), 6(187- 203), 7(221- 240), 8(275- 278), 9(286- 300),10(334- 353)
Reszty aminokwasowe przewidziane przez serwer PPM jako zanurzone w błonie	
GerAA	287-303,305-325,338-375,402-429,431,434
GerAB	13,16-36,38,41-64,74,78,82,85-87,89-137,145,148-171,174-175,187-203,206,218-252,255-256,260,275-303,306-325,331-353



Rysunek 52: Orientacja w błonie modeli kompleksów błonowej części receptora GerA. A) Model zbudowany na podstawie monomerów podjednostek A i B zorientowany w błonie przy użyciu narzędzia QMbrane. B) Orientacja kompleksu GerA zbudowanego z monomerów obliczona przy użyciu serwera PPM. Kolor niebieski oznacza wewnętrzną (tzw. cytoplazmatyczną) granicę błony, kolor czerwony – granicę zewnętrzną. C) i D) przedstawiają kompleks zbudowany z dimerów GerAA i GerAB zorientowany w błonie za pomocą QMbrane (C) oraz PPM (D).

6.6.4. Model wewnętrznej błony przetrwalnika

Model wewnętrznej błony przetrwalnika zbudowałam na podstawie informacji uzyskanych z badań prof. Banasia oraz na podstawie modelu błon biologicznych *B. subtilis*. Wyniki badań nad błonami przetrwalnika, które prowadzone były w Zakładzie Biochemii Roślin Międzyuczelnianego wydziału Biotechnologii Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego, zostały uśrednione i zebrane w tabeli 15. Wyniki te pokazują, że w przetrwalnikach najwięcej jest glikolipidów, prawdopodobnie steroli, natomiast w przetrwalnikach nieopłaszczonych (pozbawionych najbardziej zewnętrznej

warstwy) jest więcej lipidów obojętnych niż w przetrwalnikach opłaszczonych (dane niepokazane). W przetrwalnikach pozbawionych zewnętrznej warstwy znajduje się też więcej kwasów tłuszczowych o długich łańcuchach (tabela 15). Wyniki te nie uwzględniają natomiast podziału błon przetrwalników na wewnętrzne i zewnętrzne.

Badania opublikowane w 2009 roku pokazują, że skład wewnętrznej błony przetrwalnika jest istotny dla poprawnego przebiegu procesu kiełkowania oraz właściwości spor *B. subtilis* (Griffiths i Setlow, 2009). Przetrwalniki, w których składzie brakowało lipidów PE, a zawartość CL była niska wykazywały się obniżoną zdolnością kiełkowania w obecności L-alaniny oraz mieszaniny AGFK. Ponieważ zmiany w składzie błony nie wpływały na jej budowę, badacze wysnuli wniosek, iż poprawne kiełkowanie inicjowane germinantami wymaga odpowiedniego środowiska wewnątrz błony. Dlatego, skład wewnętrznej błony przetrwalników, w której umiejscowione są receptory odpowiadające za inicjację kiełkowania jest niezwykle istotnym elementem badań nad działaniem tych receptorów.

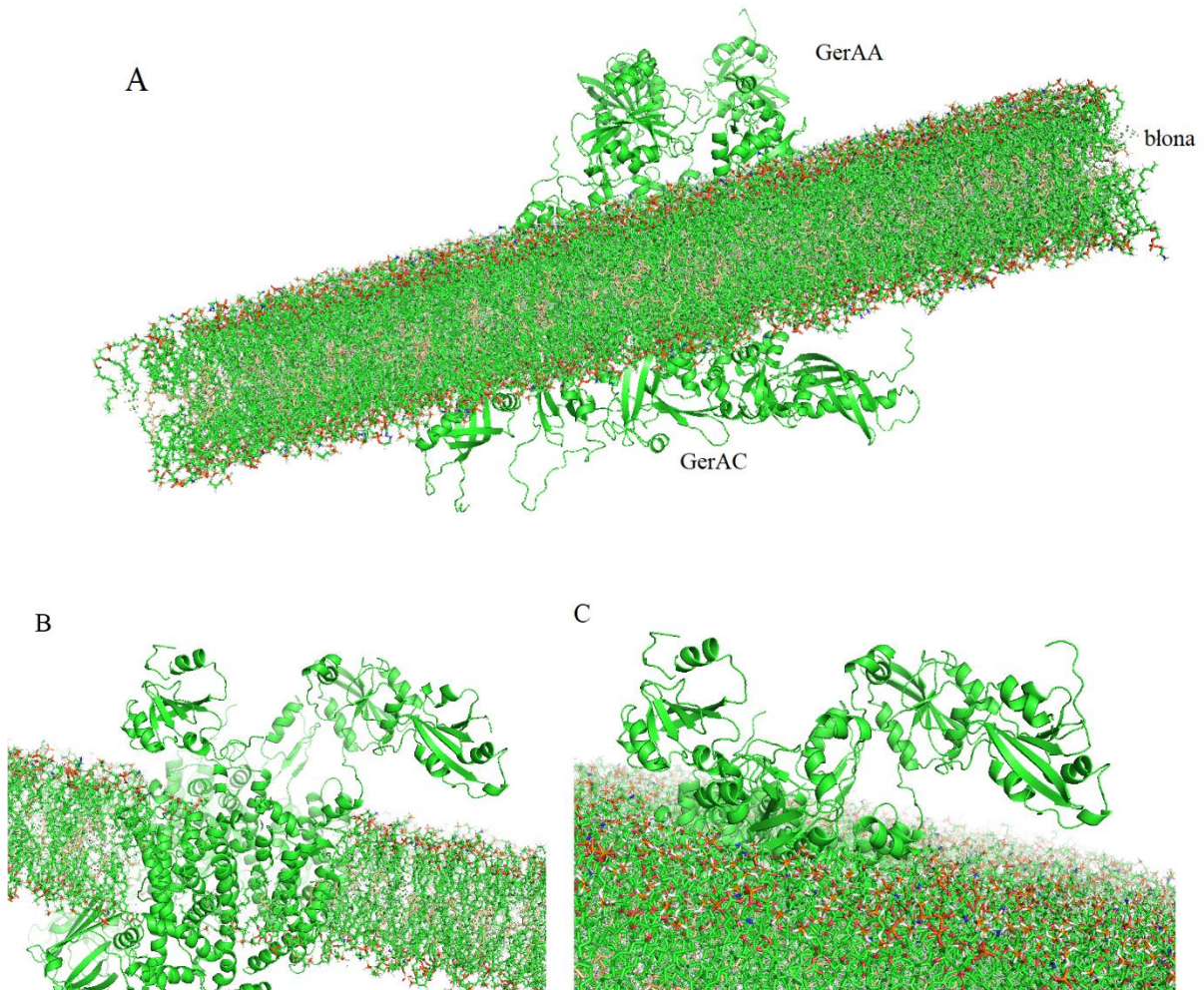
Tabela 15: Średni procentowy skład kwasów tłuszczowych opłaszczonych i nieopłaszczonych spor *Bacillus subtilis*. W przypadku spor nieopłaszczonych zaznaczone zostały znaczące różnice w ilości w porównaniu do przetrwalników opłaszczonych (↑ - więcej, ↓ - mniej)

ID	13:1	14:0	14:1	14:1'	15:1	15:1'	16:0	16:1'	16:1*	18:0	18:1	x
Opłaszczone	2,21	1,08	13,0	55,2	4,9	5,6	0,84	3,5	8,1	2,7	0,69	2,2
Nieopłaszczone	1,08	1,03	7,9↓	28,4↓	4,2	14,4↑	-	4,3	8,4	10,7↑	17,7↑	0,86

Zbudowany przeze mnie model wewnętrznej błony przetrwalnika *B. subtilis* bazuje na wyżej wymienionych wynikach, nie jest natomiast modelem wiernie oddającym skład błony, a jedynie przybliżeniem, poczynionym na potrzebę moich badań. Zbudowanie dokładnego modelu na chwilę obecną jest niemożliwe ze względu na ograniczenia narzędzi, z których korzystałam oraz na brak rozróżnienia błon na wewnętrzne i zewnętrzne w wynikach badań laboratoryjnych dotyczących składu błon przetrwalników.

Rysunek 53 przedstawia model kompleksu GerA zbudowanego z dimerów zanurzony w modelu błony fosfolipidowej. Część błonowa kompleksu, czyli dimer białek GerAB oraz fragment dimeru GerAA są całkowicie zanurzone w błonie. Ponad wewnętrzną warstwę błony wystają niebłonowe domeny białek GerAA (rysunek 53, panel C). Dimer białek GerAC niemal w całości znajduje się na zewnątrz błony. Ostateczna orientacja kompleksu w błonie bardzo przypomina obliczoną przez QMbrane i przedstawioną na rysunku 52.

Bardziej dokładny model błony pozwoliłby zbadać, za pomocą symulacji, zachowanie kompleksów w błonie. Dodatkowo można byłoby zbadać zachowanie kompleksu zanurzonego w modelach wewnętrznej błony różniącymi się składem lipidów. Wyniki zanurzenia kompleksu zbudowanego receptora GerA zbudowanego z monomerów znajdują się załączniku nr 3 w formie pliku pdb. Szerokość błony (obliczona przeze mnie za pomocą programu PyMol jako dystans między atomami) wynosi około 37Å.



Rysunek 53: Model kompleksu receptora zanurzonego w modelu błony fosfolipidowej imitującej wewnętrzną błonę przetrwalnika *B. subtilis*. A) Model pełnowymiarowy. Ponad błonę wystają białka podjednostki C (GerAC) oraz niebłonowe części dimeru białka GerAA, które dokładniej widać na panelu C. B) Model pokazujący dimer integralnego białka błonowego GerAB w błonie. C) Panel skupia się na pokazaniu niebłonowej części dimeru GerAA wystającej poza granice modelu błony.

7. DYSKUSJA

Receptor GerA kiełkowania przetrwalników *B. subtilis* jest kompleksem białek umiejscowionym w wewnętrznej błonie przetrwalnika i inicjuje kiełkowanie w obecności dwóch germinantów: L-alaniny lub L-waliny. Receptor ten zbudowany jest z trzech podjednostek, których struktury, jak dotąd, nie zostały rozwiązane za pomocą metod laboratoryjnych. W momencie rozpoczynania pracy tej badawczej skupionej nad zagadnieniem struktury receptora, dostępny był jedynie homolog podjednostki C oraz białko GerD ułatwiającego gromadzenie się receptorów w germinosom. W trakcie pracy nad modelami podjednostek pojawiały się nowe informacje strukturalne, które pozwoliły mi na podniesienie jakości modeli. Do tej pory jednak nie została poznana struktura błonowej części żadnego z białek A oraz B receptorów kiełkowania. Trudności z ustaleniem struktur spowodowane są zapewne błonowym charakterem tych białek oraz niedoskonałościami znanych metod ustalania struktur przestrzennych związanymi z pracą w środowisku błonowym.

W 2021 roku pojawiły się dwie prace badawcze opisujące modele białka GerAB oraz potencjalne kieszenie wiążące tego białka. Wcześniejsza z tych publikacji (Blinker, Vreede, Setlow i Brul, 2021) opisuje trzy modele podjednostki B zbudowane za pomocą dostępnych w sieci serwerów do modelowania na następujących szablonach: 3GI8 (dwa modele z różnych serwerów) oraz 6F34. Druga praca (Artzi i inni, 2021) wykorzystuje oprogramowanie Modeller oraz białko GkApcT (kod PDB 5OQT) jako szablon do zbudowania modelu GerAB. Obie publikacje sugerują nieco odmienne regiony jako kieszenie wiążące germinanty. Regiony te zostały przewidziane za pomocą różnych metod i w żadnym z przypadków nie wykorzystano dokowania molekularnego do sprawdzenia tych przewidywań.

Model GerAB zbudowany przeze mnie na bazie trzech szablonów — transporterów z rodziny APC — i opisany w rozdziale *Model podjednostki B* wykazuje wiele podobieństw do modeli przedstawionych w wyżej wymienionych pracach. Wszystkie modele przyjmują podobną strukturę, zbudowaną z co najmniej 12 α -helis, charakterystyczną dla białek należących do rodziny APC. Na tej podstawie mogę wnioskować, że zbudowany przeze mnie model podjednostki B receptora GerA jest modelem bliskim rzeczywistości. Można też, bazując na fakcie, że podjednostki różnych receptorów kiełkowania są do siebie homologiczne, zaryzykować stwierdzenie, że podjednostki B innych receptorów

kiełkowania również przyjmują struktury podobne do struktur białek z nadrodziny transporterów APC.

W roku 2019 opublikowana została struktura niebłonowej części (domena N-terminalna) homologa białka GerAA — podjednostki A receptora kiełkowania spor *Bacillus megaterium* (Li i inni, 2019), która umożliwiła mi poprawienie jakości mojego modelu białka GerAA, a także wymusiła powtórzenie wszystkich prowadzonych w ramach tej pracy badań dla nowego modelu. Wciąż jednak brakuje pewnych informacji o strukturze domeny błonowej, dlatego w tym przypadku musiałam posłużyć się białkiem o podobnej sekwencji oraz charakterze jako szablonie. Techniki modelowania molekularnego pozwoliły mi na utworzenie modelu białka GerAA, który osiągał dobre wyniki w algorytmach i funkcjach oceniających użytych do ewaluacji jakości modelu. Sama domena N-terminalna mojego modelu GerAA wykazuje się dużym podobieństwem do swojego homologa, przyjmując niemal identyczne pofałdowanie. Co za tym idzie, podobnie jak w przypadku GerAB, struktura ta będzie prawdopodobnie konserwowana wśród innych białek A receptorów kiełkowania przetrwalników.

Białko GerAC jest jedynym z białek receptora GerA, które nie posiada integralnej domeny błonowej, aczkolwiek pierwsze około 20 reszt aminokwasowych tego białka z końca aminowego prawdopodobnie zanurzonych jest w wewnętrznej błonie przetrwalnika lub z nią oddziałuje. Reszty aminokwasowe 3-15 w przeprowadzonych przeze mnie przewidywaniach struktury drugorzędowej przyjmują strukturę helikalną. Helisy tej nie udało mi się uzyskać w modelu struktury trzeciorzędowej ze względu na brak homologa tego fragmentu, który mógłby posłużyć za szablon. Białko GerBC, na podstawie którego zbudowałam model podjednostki C receptora GerA nie zawiera w swojej strukturze końca aminowego (Li, Setlow, Setlow i Hao, 2010). Zbudowany przeze mnie model GerAC wykazuje się dużym podobieństwem strukturalnym do swojego homologa, co czyni go wiarygodnym modelem. Zbudowanie modelu białka C receptora GerK, który przyjmuje strukturę podobną do GerBC, a zatem również GerAC, pozwala mi stwierdzić, że pofałdowanie, które obserwujemy w strukturze GerABC jest konserwowane wśród białek C receptorów kiełkowania.

Biorąc pod uwagę ocenę proponowanych przeze mnie modeli, ich jakość oraz informacje strukturalne dotyczące homologów, czy w przypadku GerAB inne modele homologiczne proponowane w literaturze, stwierdzam, że wykonane przeze mnie modele z dużym prawdopodobieństwem oddają prawdziwe struktury białek receptora GerA i stanowią dobry materiał do dalszych prac badawczych oraz do sprawdzania kolejnych

hipotez stawianych przez badaczy w temacie działania receptora GerA. Modele prawdopodobnie wymagają kilku usprawnień, których będzie można dokonać, gdy pojawią się nowe informacje dotyczące błonowych części receptora.

7.1. Model receptora GerA

Receptory kiełkowania przetrwalników są badane od wielu dekad, jednak do tej pory nie poznano ich budowy, poza niebłonowymi fragmentami struktur, czy strukturami pojedynczych podjednostek. Nie znamy też mechanizmu oddziaływania receptora z ligandem ani tego, jak receptor oddziałuje z błoną oraz jak podjednostki wiążą się ze sobą w kompleks receptora. Uważa się, że stechiometria białek receptora wynosi 1:1:1 i właśnie ta informacja była bazą, na podstawie której budowałam modele receptora zaprezentowane w tej pracy badawczej.

Na podstawie zbudowanych przeze mnie modeli podjednostek zaproponowałam dwa modele receptora GerA, inicjującego kiełkowanie spor w obecności L-alaniny i/lub L-waliny. Pierwszy model receptora bazuje na podobieństwach procesu kiełkowania do procesów chemotaksji bakteryjnej i zakłada, że receptor GerA, zbudowany z trzech podjednostek, przyjmuje formę trimeru dimerów. Taki model receptora jest niezwykle skomplikowany i stabilizowany siecią licznych oddziaływań pomiędzy podjednostkami. Być może ułożenie podjednostek w rzeczywistości nieznacznie różni się od zbudowanego przeze mnie modelu teoretycznego, ponieważ orientacja w błonie tego modelu była niezwykle trudna do ustalenia. Pomimo tych trudności, prezentowany przeze mnie model zapewnia dużo informacji strukturalnych, na bazie których można kontynuować badania nad strukturą receptora. Przykładowo, poprzez mutagenezę miejsc wiązania podjednostek do siebie.

Dodatkowym argumentem za przyjęciem trimeru dimerów jako prawdopodobnego modelu jest zachowanie szablonów białek GerAB oraz GerAC. Szablony (transportery z rodziny APC), na podstawie których zbudowany został model GerAB łączą się w dimery i właśnie w formie dimerycznej wiążą ligand (Shaffer, Goehring, Shankaranarayanan i Gouaux, 2009). Homolog GerAC — białko GerBC — z kolei dimeryzuje w kryształach (Li, Setlow, Setlow i Hao, 2010), co pozwala sądzić, że przy dużym stężeniu białek, a o takim prawdopodobnie mówimy w przypadku germinosomu, GerAC będzie dążyć do dimeryzacji.

Pomimo wielu argumentów przemawiających za modelem GerA opartym o strukturę trimeru dimerów, model receptora, który zbudowałam z monomerów wydaje się być równie obiecującym obiektem dalszych badań. Głównie ze względu na brak badań opisujących dimeryczny charakter podjednostek receptora w literaturze.

Pełny model GerA zbudowany z dimerów wymaga pewnej korekty, która umieściłaby dimer GerAC całkowicie poza błoną. Prawdopodobnie struktura dimeru GerAC proponowana przeze mnie w tym modelu nie jest w pełni optymalna ze względu na brak 23 N-końcowych reszt aminokwasowych.

Wydaje się, że poprawienie modelu GerAA w zakresie domeny błonowej mogłoby ułatwić zanurzenie tego białka w błonie, a tym samym poprawić jakość modelu receptora GerA. W chwili obecnej w literaturze nie ma struktury bliskiego homologa, która mogłaby posłużyć za szablon dla domeny błonowej. Transportery bakteryjne użyte przeze mnie w toku prac nad modelem GerAA są wystarczająco podobne na poziomie sekwencji, by na ich podstawie stworzyć model struktury. Proponowany przeze mnie model wspierany informacjami uzyskanymi z przewidywań fragmentów transbłonowych jest oceniany wysoko i wystarczająco dokładny, by proponować model receptora czy badać miejsca wiązania germinantów, niemniej posiadane na tym etapie bardziej dokładnych danych strukturalnych powiązanych z częścią błonową białka wydaje się niezbędne do przeprowadzenia kolejnych badań nad zachowaniem receptora GerA w błonie. Kolejnym obszarem, któremu należałoby poświęcić więcej badań jest skład wewnętrznej błony przetrwalnika. Błony biologiczne *B. subtilis* wciąż są szeroko badane przez różne grupy badawcze (najnowsze doniesienia w tym temacie pochodzą z grudnia 2021 (van Tilburg, Warmer, van Heel, Sauer i Kuipers, 2021)), jednak wciąż brakuje konkretnych informacji ilościowych, które pomogłyby by w dokładnym odtworzeniu środowiska receptora GerA. Jak wiemy z przeprowadzonych badań, skład wewnętrznej błony jest istotny dla poprawnego oddziaływania receptora z błoną, a także wpływa na proces kiełkowania (Griffiths i Setlow, 2009).

7.2. Wiązanie germinantów

Jak pokazują rysunki opisujące wyniki przeprowadzonych przeze mnie dokowań dla GerAA i GerAB oraz ich dimerów, potencjalne kieszenie wiążące wytypowane przez eksperymenty dokowania umiejscowione są w błonowych częściach białka GerAB,

błonowej części GerAA oraz części niebłonowej GerAA, znajdującej się wewnątrz przetrwalnika. Umieszczenie kieszeni wiążącej w domenach receptora znajdujących się w rdzeniu przetrwalnika może wiązać się z tym, że germinant (L-alanina i/lub L-walina) będzie musiał w pierwszej kolejności spenetrować błonę, gdyż dopiero wtedy będzie mógł połączyć się z receptorem i rozpocząć proces kiełkowania. Możliwym jest, patrząc na uzyskane wyniki dokowań do podjednostek i modelu kompleksu, że wiązanie germinantów do receptora GerA może mieć charakter tymczasowy, a ligand być transportowany przez jedną podjednostkę do drugiej, co może sprawiać, że wędruje on w głąb błony, wiążąc się po drodze do kilku miejsc wiążących obu podjednostek receptora. Takie przekazywanie germinanta może być niezbędne do aktywacji receptora, co tłumaczyłoby wyniki badań pokazujące różne kieszenie wiążące białka GerAB oraz opisujące liczne mutacje w różnych częściach podjednostek A i B, których występowanie zaburza poprawne działanie receptora. Może to świadczyć o tym, że L-alanina lub L-walina, po przyłączeniu się do receptora, poprzez swoją wędrówkę pomiędzy podjednostkami, aktywują receptor (przykładowo poprzez zmiany konformacyjne odwiedzonych przez ligand kieszeni wiążących), a tym samym rozpoczynają inicjację kiełkowania. Ta hipoteza wymaga dokładnego zbadania ścieżki liganda w kompleksie oraz licznych badań *in vivo*, które mogłyby potwierdzić lub obalić to założenie. Na tym etapie wiemy natomiast, że liczne mutacje w białkach A i B, które umiejscowione są w różnych regionach białek, wpływają na poprawne funkcjonowanie receptora (Moir, Lafferty i Smith, 1979; Mongkoltharuk, Cooper, Mawer, Allan i Moir, 2011).

Tabela 13 opisuje najczęściej występujące regiony w wynikach dokowań globalnych germinantów. Ze względu na częstotliwość wiązania się L-alaniny oraz L-waliny do tych miejsc, mogą uznać wymienione regiony za potencjalne kieszenie wiążące. Regiony te występują zarówno w białku GerAB jak i GerAA. W badaniach laboratoryjnych przeprowadzonych w zakładzie Bakteriologii Molekularnej Międzyuczelnianego Wydziału Biotechnologii UG i GUMed, pokazaliśmy, że reszty aminokwasowe wymienione w regionie pierwszym oraz znajdujące się w ich sąsiedztwie są istotne dla poprawnej odpowiedzi receptora na L-alaninę (reszty 299, 303 oraz 347 białka GerAA).

8. BIBLIOGRAFIA

- [1] Andoh, Y., Hayakawa, S. i Okazaki, S. (2020). Molecular dynamics study of lipid bilayers modeling outer and inner leaflets of plasma membranes of mouse hepatocytes. I. Differences in physicochemical properties between the two leaflets. *The Journal of Chemical Physics*, 153(3).
- [2] Andre, I., Bradley, P., Wang, C. i Baker, D. (2007). Prediction of the structure of symmetrical protein assemblies. *Proc Natl Acad Sci*, 104(45), strony 17656-17661.
- [3] Anfinsen, C. (1973). Principles that govern the folding of protein chains. *Science*, 181(4096), strony 223-230.
- [4] Artzi, L., Alon, A., Brock, K., Green, A., Tam, A., Ramirez-Guadiana, F., . . . Rudner, D. (2021). Dormant spores sense amino acids through the B subunit of their germination receptors. *Nature Communications*, 12(1).
- [5] Bajaj, M. i Blundell, T. (1984). Evolution and tertiary structure of proteins. *Annual Review of Biophysics and Bioengineering*, 13(1), strony 453-492.
- [6] Berg, J. M., Tymoczko, J. L. i Stryer, L. (2005). *Biochemia*. W J. M. Berg, J. L. Tymoczko i L. Stryer. Warszawa: PWN.
- [7] Bishop, D., Rutberg, L. i Samuelsson, B. (1967). The chemical composition of the cytoplasmic membrane of *Bacillus subtilis*. *European Journal of Biochemistry*, 2(4), strony 448-453.
- [8] Blinker, S., Vreede, J., Setlow, P. i Brul, S. (2021). Predicting the structure and dynamics of membrane protein GerAB from *Bacillus subtilis*. *International Journal of Molecular Sciences*, 22(7), str. 3793.
- [9] Cavanagh, J., Fairbrother, W. i Palmer, A. (2007). *Protein NMR Spectroscopy: Principles and Practice*. Londyn: Elsevier Academic Press.
- [10] Cheng, J. (2008). A multi-template combination algorithm for protein comparative modeling. *BMC Structural Biology*, 8(18).
- [11] Clejan, S., Krulwich, T., Mondrus, K. i Seto-Toung, D. (1986). Membrane lipid composition of obligately and facultatively alkalophilic strain of *Bacillus* spp. *Journal of Bacteriology*, 168(1), strony 334-340.
- [12] Cohn, F. (1872). Untersuchungen über Bakterien. *Beitr Biol Pflanz*, 1, strony 27-224.
- [13] Cooper, G. i Moir, A. (2011). Amino acid residues in the GerAB proteins important in the function and assembly of the alanine spore germination receptor of *Bacillus subtilis*. *Journal of Bacteriology*, 193(3), strony 2261-2267.
- [14] Cowan, A., Koppel, D., Setlow, B. i Setlow, P. (2003). A soluble protein is immobile in dormant spores of *Bacillus subtilis* but is mobile in germinated spores: Implications for spore dormancy. *PNAS*, 100(7), strony 4209-4214.
- [15] Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunstere, W. i Mark, A. (2004). Peptide Folding : When Simulation Meets Experiment. *Angewandte Chemie*, 38(1-2), strony 236-240.
- [16] Den Kamp, J., Redai, I. i Van Deenen, L. (1996). Phospholipid composition of *Bacillus subtilis*. *Journal of Bacteriology*, 99(1), strony 298-303.
- [17] Desta, I., Porter, K., Xia, B., Kozakov, D. i Vajda, S. (2020, Wrzesień). Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure*, 28(9), strony 1071-1081.
- [18] Drozdetskiy, A., Cole, C., Procter, J. i Barton, G. (2015, Kwiecień 16). *Nucl. Acids Res*.
- [19] Elokely, K. i Doerksen, R. (2013). Docking Challenge: Protein Sampling and Molecular Docking Performance. *Journal of Chemical Information and Modeling*, 53(8), strony 1934-1945.
- [20] Errington, J. (1993). *Bacillus subtilis* sporulation: Regulation of gene expression and control of morphogenesis. *Microbiological Reviews*, 57(1), strony 1-33.
- [21] Errington, J. (2003). Regulation of endospore formation in *Bacillus subtilis*. *Nature Reviews Microbiology*, 1(2), strony 117-126.
- [22] Falke, J. (2002). Cooperativity between bacterial chemotaxis receptors. *PNAS*, 99(10), strony 6530-6532.
- [23] Feenstra, K., Abeln, S. i Heringa, J. (2018). Protein Three-Dimensional Structure Prediction. W *Encyclopedia of Bioinformatics and Computational Biology*. Elsevier.
- [24] Fiser, A. (2010). Template-Based Protein Structure Modeling. W *Computational Biology. Methods in Molecular Biology*. (strony 73-94). Humana Press, Totowa, NJ.
- [25] Gao, X., Zhou, L., Jiao, X., Lu, F., Yan, C., Zeng, X., . . . Shi, Y. (2010). Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature*, 463(7282), strony 828-832.
- [26] Gould, G. i Sale, J. (1970). Initiation of bacterial spores by hydrostatic pressure. *Journal of General Microbiology*, 60(3), strony 335-346.
- [27] Greer, J. (1981). Comparative model-building of the mammalian serine proteases. *Journal of Molecular Biology*, 153(4), strony 1027-1042.
- [28] Greła, A. (2019). Warianty receptora GerA w przetrwalnikach różnych szczepów laboratoryjnych *Bacillus subtilis*. Gdańsk.

- [29] Grela, A., Jamrozek, I., Hubisz, M., Iwanicki, A., Hinc, K., Kazmierkiewicz, R. i Obuchowski, M. (2018). Positions 299 and 302 of the GerAA subunit are important for function of the GerA spore germination receptor in *Bacillus subtilis*. *PLOS one*, 13(6).
- [30] Griffiths, K. i Setlow, P. (2009). Effects of modification of membrane lipid composition on *Bacillus subtilis* sporulation and spore properties. *Journal of Applied Microbiology*, 106(6), strony 2064-2078.
- [31] Griffiths, K., Zhang, J., Cowan, A., Yu, J. i Setlow, P. (2011). Germination proteins in the inner membrane of dormant *Bacillus subtilis* spores colocalize in a discrete cluster. *Molecular Microbiology*, 81(4), strony 1061-1077.
- [32] Hofmann, K. i Stoffel, W. (1993). TMbase - a database of membrane spanning proteins segments. *Biol. Chem.*
- [33] Holm, L. (2020). Using Dali for Protein Structure Comparison. W G. Z., *Structural Bioinformatics. Methods in Molecular Biology* (strony 29-42). New York: Humana.
- [34] Holm, L. i Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends in Biochemical Sciences*, 20(11), strony 478-480.
- [35] Huang, S.-Y. i Zou, X. (2010). Advances and Challenges in Protein-Ligand Docking. *International journal of molecular science*, 11(8), strony 3016-3034.
- [36] Hudson, K., Corfe, B., Kemp, E., Feavers, I., Coote, P. i Moir, A. (2001). A Localization of GerAA and GerAC germination proteins in *Bacillus subtilis* spore. *J Bacteriol.*, 183(14), strony 4317-4322.
- [37] Igarashi, T. i Setlow, P. (2005). Interaction between individual protein components of the GerA and GerB nutrient receptors that trigger germination of *Bacillus subtilis* spores. *Journal of Bacteriology*, 187(7), strony 2513-2518.
- [38] Illergård, K., Ardell, D. H. i Elofs, A. (2009). Structure is three to ten times more conserved than sequence--a study of structural response in protein cores. *Proteins*, 77(3), strony 499-508.
- [39] Jabeen, A., Mohamedali, A. i Ranganathan, S. (2018). Protocol for Protein Structure Modelling. W *Encyclopedia of Bioinformatics and Computational Biology*. Elsevier Inc.
- [40] Jack, D., Paulsen, I. i Saier, M. (2000). The amino acid/polyamine/organocation APC superfamily of transporters specific for amino acids, polyamines and organocations. *Microbiology*, 146(8), strony 1797-1814.
- [41] Jiang, D., Zhao, Y., Wang, X., Fan, J., Heng, J., Xuehui, L., . . . Zhang, X. (2013). Structure of the YajR transporter suggests a transport mechanism based on the conserved motif A. *PNAS*, 110(36), strony 14664-14669.
- [42] Jo, S., Kim, T. i Im, W. (2007). Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLOS ONE*, 2(9).
- [43] Jones, D., Taylor, W. i Thornton, J. (1992). A new approach to protein fold recognition. *Nature*, 358(6381), strony 86-89.
- [44] Jones, S. i Thornton, J. (1996). Principles of protein-protein interactions. *Proc. Natl. Acas. Sci.*, 93(1), strony 13-20.
- [45] Kawai, F., Shoda, M., Harashima, R., Sadaie, Y., Hara, H. i Matsumoto, K. (2004). Cardiolipin domains in *Bacillus subtilis* Margburg membranes. *Journal of Bacteriology*, 186(5), strony 1475-1483.
- [46] Kelley, L. i Sternberg, M. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, 4(3), strony 363-371.
- [47] Kendrew, J., Bodo, G., Dintzis, H., Parrish, R., Wyckoff, H. i Phillips, D. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*, 181(4610), 662-666.
- [48] Kentner, D. i Sourjik, V. (2006). Spatial organization of the bacterial chemotaxis system. *Current Opinion in Microbiology*, 9(6), strony 619-624.
- [49] Keskin, O., Tuncbag, N. i Gursoy, A. (2016). Predicting protein-protein interactions from the molecular to the proteome level. *Chemical Reviews*, 116(8), strony 4884-4909.
- [50] Khor, B., Tye, G., Lim, T. i Choong, Y. (2015). General overview on structure prediction of twilight-zone proteins. *Theoretical Biology and Medical Modelling*, 12(14).
- [51] Kim, S.-H., Wang, W. i Kim, K. (2002). Dynamic and clustering model of bacterial chemotaxis receptors: Structural basis for signaling and high sensitivity. *PNAS*, 99(18), strony 11611-11615.
- [52] Korza, G. i Setlow, P. (2013). Topology and accessibility of germination proteins in the *Bacillus subtilis* spore inner membrane. *J Bacteriol.*, 195(7), strony 1484-1491.
- [53] Kozakov, D., Beglov, D., Bohuund, T., Mottarella, S., Xia, B., Hall, D. i Vajda, S. (2013). How good is automated protein docking? *Proteins: Structure, Function and Bioinformatics*, 81(12), strony 2159-2166.
- [54] Kozakov, D., Hall, D., Xia, B., Porter, K., Padhomy, D., Yueh, C., . . . Vajda, S. (2017). The ClusPro web server for protein-protein docking. *Nature Protocols*, 12(2), strony 255-278.
- [55] Kozlovskii, I. i Popov, P. (2020). Spatiotemporal identification of druggable binding sites using deep learning. *Communications biology*, 3(1), strony 1-12.
- [56] Kryshchovych, A., Barbato, A., Monastyrsk, B., Fidelis, K., Schwede, T. i Tramontano, A. (2016). Methods of model accuracy estimation can help selecting the best models from decoy sets:

- Assessment of model accuracy estimations in CASP11. *Proteins: Structure, Function, and Bioinformatics*, S1(84), strony S349–S369.
- [57] Kufareva, I. i Abagyan, R. (2012). Methods of protein structure comparison. *Methods in Molecular Biology*, 857, strony 231–257.
- [58] Kufareva, I., Handel, T. i Abagyan, R. (2015). Experiment-Guided Molecular Modeling of Protein-Protein Complexes Involving GPCRs. *Methods in Molecular Biology*, 1335, strony 295–311.
- [59] Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azavedo, V., . . . Borchert, S. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature*, 390(6657), str. 249.
- [60] Kuntz, I., Blaney, J., Oatley, S., Langridge, R. i Ferrin, T. (1982). A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*, 161(2), strony 269-288.
- [61] Larsson, P., Wallner, B., Lindahl, E. i Elofsson, A. (2008). Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Science*, 17(6), strony 990-1002.
- [62] Laskowski, R. i Swindells, M. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling*, 51(10), strony 2778-2786.
- [63] Lee, J., Freddolino, P. i Zhang, Y. (2017). Ab Initio Protein Structure Prediction. W D. Rigden, *From Protein Structure to Function with Bioinformatics*. Dordrecht: Springer.
- [64] Lei, H., Chou, T., Su, C., Bolla, J., Kumar, N., Radhakrishnan, A., . . . Yu, E. (2014). Crystal structure of the open state of the *Neisseria gonorrhoeae* MtrE outer membrane channel. *PLoS One*, 9(6).
- [65] Li, Y., Jin, K., Ghosh, S., Devarakonda, P., Carlson, K., Davis, A., . . . Hao, B. (2014). Structural and functional analysis of the GerD spore germination protein of *Bacillus* species. *Journal of Molecular Biology*, 426(9), strony 1995-2008.
- [66] Li, Y., Jin, K., Perez-Valdespino, A., Federkiewicz, K., Davis, A., Maciejewski, M., . . . Hao, B. (2019). Structural and functional analyses of the N-terminal domain of the A subunit of a *Bacillus megaterium* spore germinant receptor. *PNAS*, 116(23), strony 11470-11479.
- [67] Li, Y., Setlow, B., Setlow, P. i Hao, B. (2010). Crystal Structure of the GerBC Component of *Bacillus Subtilis* Spore Germinant Receptor. *Journal of Molecular Biology*, 402(1), str. 8016.
- [68] Lomize, M., Pogozheva, I., Joo, H., Mosberg, H. i Lomize, A. (2012). OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Research*, 40(Database issue), strony 370-376.
- [69] Long, F., Su, C., M.T., Boyken, S., Rajashankar, K., Jernigan, R. i Yu, E. (2010). Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. *Nature*, 467(7314), strony 484-488.
- [70] Marrink, S., Corradi, V., Souza, P. C., Ingolfsson, H. I., Tielman, D. i Sansom, M. S. (2019). Computational modeling of realistic cell membranes. *Chemical Reviews*, 119(9), strony 6184-6226.
- [71] Marti-Renom, M., Capriotti, E., Shindyalov, I. i Bourne, P. (2009). Structure comparison and alignment. W J. Gu i P. Bourne, *Structural Bioinformatics* (strony 397–418). John Wiley & Sons, Inc.
- [72] Marti-Renom, M., Stuart, A., Fiser, A., Melo, F. i Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.*, 29, strony 291-325.
- [73] Mccammon, J. A., Gelin, B. R. i Karplus, M. (1977). Dynamics of proteins. *Nature*, 267(5612), strony 585-590.
- [74] Moir, A. (2006). How do pores germinate? *Journal of Applied Microbiology*, 101(3), strony 526-530.
- [75] Moir, A., Kemp, E., Robinson, C. i Corfe, B. (1994). The genetic analysis of bacterial spore germination. *Journal of Applied Microbiology*, 77(3), strony 9S-14S.
- [76] Moir, A., Lafferty, E. i Smith, D. (1979). Genetic analysis of spore germination mutants of *Bacillus subtilis* 168: the correlation of phenotype with map location. *Journal of General Microbiology*, 111(1), strony 165-180.
- [77] Mongkolthanaruk, W., Cooper, G., Mawer, J., Allan, R. i Moir, A. (2011, May). Effect of Amino Acid Substitutions in the GerAA Protein on the Function of the Alanine-Responsive Germinant Receptor of *Bacillus subtilis* Spores. *Journal of Bacteriology*, 193(9), strony 2268-2275.
- [78] MSA. (brak daty). Pobrano z lokalizacji <https://www.ebi.ac.uk/Tools/msa/>
- [79] Nakamura, T., Oda, T., Fukasawa, Y. i Tomii, K. (2017). Template-based quaternary structure prediction of proteins using enhanced profile-profile alignments. *Proteins*, 86 Suppl 1(Suppl Suppl 1), strony 274-282.
- [80] Nickels, J., Chatterjee, S., Mostofian, B., Stanley, C., Ohl, M., Zolnierczuk, P., . . . Katsaras, J. (2017). *Bacillus subtilis* lipid extract, a branched-chain fatty acid model membrane. *The Journal of Physical Chemistry Letters*, 8(17), strony 4214-4217.
- [81] Nickles, J. D., Chatterjee, S., Stanley, C., Qian, S., Cheng, X., Myles, D., . . . Katsaras, J. (2017). The in vivo structure of biological membranes and evidence for lipid domains. *PLoS Biology*, 15(5).
- [82] Nooren, I. i Thornton, J. (2003). Diversity of protein-protein interactions. *The EMBO Journal*, 22(14), strony 3486-3492.

- [83] Novikov, F. i Chilov, G. (2009). Molecular docking: theoretical background, practical applications and perspectives. *Mendeleev Communications*, 19(5), strony 237–242.
- [84] Pace, C. i Scholtz, J. (1998). A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophysical Journal*, 75(1), strony 422–427.
- [85] Paidhungat, M. i Setlow, P. (1999). Isolation and characterization of mutations in *Bacillus subtilis* that allow spore germination in the novel germinant D-alanine. *Journal of Bacteriology*, 181(11), strony 3341–3350.
- [86] Paidhungat, M. i Setlow, P. (2000). Role of Ger proteins in nutrient and nonnutrient triggering of spore germination in *Bacillus subtilis*. *Journal of Bacteriology*, 182(9), strony 2513-2519.
- [87] Paidhungat, M. i Setlow, P. (2002). Spore germination and outgrowth. W J. Hoch, R. Losick i A. Sonenshein, *Bacillus subtilis and its Relatives: From Genes to Cells* (strony 537-548). Washington DC: American Society for Microbiology.
- [88] Paidhungat, M., Setlow, B., Daniels, W., Hoover, D., Papafragkou, E. i Setlow, P. (2002). Mechanisms of Induction of Germination of *Bacillus subtilis* Spores by High Pressure. *Applied and Environmental Microbiology*, 68(6), strony 3172-3175.
- [89] Pelczar, P. i Setlow, P. (2008). Localization of the Germination Protein GerD to the Inner Membrane in *Bacillus subtilis* spores. *Journal of Bacteriology*, 190(16), strony 5635–5641.
- [90] Pelczar, P., Igarashi, T., Setlow, B. i Setlow, P. (2007). Role of GerD in Germination of *Bacillus subtilis* Spores. *Journal of Bacteriology*, 189(3), strony 1090-1098.
- [91] Persson, B. i Argos, P. (1994). Prediction of transmembrane segments in proteins utilizing multiple sequence alignments. *Journal of Molecular Biology*, 237(2), strony 182-192.
- [92] Persson, B. i Argos, P. (1996). Topology prediction of membrane proteins. *Protein Science*, 5(2), strony 363-371.
- [93] Pettersen, E., Goddard, T., Huang, C., Couch, G., Greenblatt, D., Meng, E. i Ferrin, T. (2004). UCSF Chimera - a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), strony 1605-1612.
- [94] Powell, J. i Strange, R. (1953). Biochemical changes occurring during the germination of bacterial spores. *The Biochemical Journal*, 54(2), strony 205-209.
- [95] Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., . . . van der Spoel, D. (2013). GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7), strony 845-854.
- [96] PSA. (brak daty). Pobrano z lokalizacji <https://www.ebi.ac.uk/Tools/psa/>
- [97] PSVS. (brak daty). Pobrano z lokalizacji <https://montelionelab.chem.rpi.edu/PSVS/>
- [98] Qin, S. i Zhou, H.-X. (2007). meta-PPISP: a meta web server for protein-protein interaction site prediction. *Bioinformatics*, 23(24), strony 3386-3387.
- [99] QMEAN. (2020, 12 30). Pobrano z lokalizacji <https://swissmodel.expasy.org/qmean/>
- [100] Ramachandran, G., Ramakrishnan, C. i Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, 7, strony 95-99.
- [101] Read, R. i Chavali, G. (2007). Assessment of CASP7 predictions in the high accuracy template-based modeling category. *Proteins: Structure, Function, and Bioinformatics*, 69(S8), strony S27–S37.
- [102] Ritchie, D. i Grudinin, S. (2006). Spherical Polar Fourier Assembly of Protein Complexes with Arbitrary Point Group Symmetry. *Journal of Applied Crystallography*, 49(1), strony 158-167.
- [103] Roi, A., Kucukural, A. i Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5(4), strony 725-738.
- [104] Salentin, S., Schreiber, S., Haupt, V., Adasme, M. i Schroeder, M. (2015). PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Research Advance*, 43(1), strony W443-W447.
- [105] Sali, A. i Blundell, T. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), strony 779-815.
- [106] Sammons, R., Moir, A. i Smith, D. (1981). Isolation and properties of spore germination mutants of *Bacillus subtilis*-168 deficient in the initiation of germination. *Journal of General Microbiology*, 124(2), strony 229-241.
- [107] Sanner, M. (1999). Python: A Programming Language for Software Integration and Development. *Journal of Molecular Graphics & Modelling*, 17(1), strony 57-61.
- [108] Schlick, T. (2010). *Molecular Modeling and Simulation: an interdisciplinary guide* (wyd. 2). Nowy Jork: Springer.
- [109] Schrödinger, L. (2015). The PyMOL Molecular Graphics System, Version 2.0.
- [110] Setlow, B., Wahome, P. i Setlow, P. (2008). Release of Small Molecules during Germination of Spores of *Bacillus Species*. *Journal of Bacteriology*, 190(3), strony 4759–4763.
- [111] Setlow, P. (2006). Spore germination. *Current Opinion in Microbiology*, 6(6), strony 550-556.

- [112] Setlow, P. (2006). Spores of *Bacillus subtilis*: their resistance to radiation, heat and chemicals. *Journal of Applied Microbiology*, 101(3), strony 514-525.
- [113] Setlow, P. (2014). Germination of spores of *Bacillus* species: what we know and do not know. *Journal of Bacteriology*, 169(7), strony 1297-1305.
- [114] Setlow, P. i Johnson, E. (2012). Spores and their significance. W B. R. Doyle M.P., *Food microbiology: fundamentals and frontiers* (strony 45-79). Washington DC: ASM Press.
- [115] Shaffer, P., Goehring, A., Shankaranarayanan, A. i Gouaux, E. (2009). Structure and mechanism of a Na⁺-independent amino acid transporter. *Science*, 325(5943), strony 1010-1014.
- [116] Sherman, M. (1982). Hypothesis: Interaction between taxis and sporulation systems. *FEBS LETTERS*, 148(2), strony 192-197.
- [117] Shrout, A., Montefusco, D. i Weis, R. (2003). Template-Directed Assembly of Receptor Signaling Complexes. *Biochemistry*, 42(46), strony 13379-13385.
- [118] Sonnhammer, E. L., von Heijne, G. i Krogh, A. (1998). A hidden Markov model for prediction transmembrane helices in protein sequences. W E. J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff i C. Sensen, *In Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology* (strony 175-182). Menlo Park: AAAI Press.
- [119] Sottriffer, C. i Klebe, G. (2002). Identification and mapping of small-molecule binding sites in proteins: computational tools for structure-based drug design. *Il farmaco (Societa Chimica Italiana : 1989)*, 57(3), strony 243-251.
- [120] Sourjik, V. i Armitage, J. (2010). Spatial organization in bacterial chemotaxis. *The EMBO Journal*, 29(16), strony 2274-2733.
- [121] Steward, K.-A. i Setlow, P. (2013). Numbers of individual nutrient germinant receptor and other germination proteins in spores of *Bacillus subtilis*. *Journal of Bacteriology*, 195(16), strony 3575-3582.
- [122] Studer, G., Biasini, M. i Schwede, T. (2014). Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics*, 30(17), strony i505-i511.
- [123] Studer, G., Biasini, M. i Schwede, T. (2014). Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics*, 30(17), strony i505-i511.
- [124] Studer, G., Rempfer, C., Waterhouse, A., Gumienny, G., Haas, J. i Schwede, T. (2020). QMEANDisCo - distance constraints applied on model quality estimation. *Bioinformatics*(36), strony 1765-1771.
- [125] Studer, G., Rempfer, C., Waterhouse, A., Gumienny, G., Haas, J. i Schwede, T. (2020). QMEANDisCo - distance constraints applied on model quality estimation. *Bioinformatics*, 36(6), strony 1765-1771.
- [126] Su, C., Long, F., Lei, H., Bolla, J., Do, S., Rajashankar, K. i Yu, E. (2012). Charged Amino Acids (R83, E567, D617, E625, R669, and K678) of CusA Are Required for Metal Ion Transport in the Cus Efflux System. *Journal of Molecular Biology*, 422(3), strony 429-441.
- [127] Ślósarek, G. (2011). *Biofizyka Molekularna*. Warszawa: Wydawnictwo Naukowe PWN.
- [128] Tharckay, P., Behravan, T., Southworth, T. i Moir, A. (2001). GerN, an antiporter homologue important i the germination of *Bacillus cereus* endospores. *Journal of Bacteriology*, 183(2), strony 746-482.
- [129] Tovchigrechko, A. i Vakser, I. (2006). GRAMM-X public erb server for protein-protein docking. *Nucleic Acids Research*, 34(suppl_2), strony W310-W314.
- [130] Trott, O. i Olson, A. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, 31(2), strony 455-461.
- [131] Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S., Xia, B., . . . Kozakov, D. (2017). New additions to the ClusPro server motivated by CAPRI. *Proteins: Structure, Function and Bioinformatics*, 85(3), strony 435-444.
- [132] van Meer, G., Voelker, D. i Feigenson, G. (2008). Membrane lipids: where they are and how they behave. *Nature reviews: Mocolular Cell Biology*, 9(2), strony 112-124.
- [133] van Tilburg, A., Warmer, F., van Heel, A., Sauer, U. i Kuipers, O. (2021). Membrane composition and organization of *Bacillus subtilis* 168 and its genome-reduced derivative miniBacillus PG10. *Microbial Biotechnology*, 15(5), strony 1633-1651.
- [134] Venclovas, C. i Margelevicius, M. (2005). Comparative modeling in CASP6 using consensus approach to template selection, sequence-structure alignment, and structure assessment. *Proteins*, 61(7), strony 99-105.
- [135] Voet, D., Voet, J. i Pratt, C. (2016). *Fundamentals of biochemistry: Life at th emolecular level* (wyd. 5). ed.Wiley.

- [136] Vreeland, R., Rosenzweig, D. i Powers, D. (2000). Isolation of a 250 million-year-old halotolerant bacterium from primary salt crystal. *Nature*, 407(6806), strony 897-900.
- [137] Webb, B. i Sali, A. (2016). Comparative Protein Structure Modeling Using Modeller Chapter 5:Unit-5.6. *Current Protocols in Bioinformatics*.
- [138] Wilson, M., Carlson, P., Janes, B. i Hanna, P. (2012). Membrane topology of *Bacillus anthracis* GerH germinant receptor proteins. *Journal of Bacteriology*, 194(6), strony 1369-1377.
- [139] Wu, E., Cheng, X., Jo, S., Rui, H., Song, K., Dávila-Contreras, E., . . . Im, W. (2014). CHARMM-GUI Membrane Builder Toward Realistic Biological Membrane Simulations. *Journal of Computational Chemistry*, 35(27), strony 1997-2014.
- [140] Wuytack, E., Soons, J., Poschet, F. i Michiels, C. (2000). Comparative Study of Pressure- and Nutrient-Induced Germination of *Bacillus subtilis* spores. *Applied and Environmental Microbiology*, 66(1), strony 257-261.
- [141] Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. i Zhang, Y. (2015). The I-TASSERSuite: Protein structure and function prediction. *Nature Methods*, 12(1), strony 7-8.
- [142] Yuanpeng, J., Powers, R. i Montelione, G. (2005). Protein NMR Recall, Precision, and F-measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics. *Journal of the American Chemical Society*, 127(6), strony 1665-1674.
- [143] Zhang, Y. (2008). I-TASSER server for protein 3D prediction. *BMC Bioinformatics*, 9(40).
- [144] Zhao, J., Cao, Y. i Zhang, L. (2020). Exploring the computational methods for protein-ligand binding site prediction. *Computational and Structural Biotechnology Journal*, 18, strony 417-426.
- [145] Zhou, H.-X. i Shan, Y. (2001). Prediction of protein interaction sites from sequence profiles and residue neighbor list. *Proteins*, 44(3), strony 336-343.
- [146] Berggard, T.; Linse, S.; James, P. (2007). Methods for the Detection and Analysis of Protein-Protein Interactions. 7(16), 2833– 2842.

9. ZAŁĄCZNIKI

Załącznik nr 1

Skrypt użyty do wielokrotnego uruchomienia programu AutoDock Vina:

```
#!/bin/bash
x=1
while [ $x -le 2250 ]
do
vina --config conf.txt --out all$x.pdbqt --log
log$x.txt>smieci$x
    x=$(( $x + 1 ))
done
```


Załącznik nr 2

Dorobek naukowy autorki pracy

Publikacje:

Grela A, **Jamrozek I**, Hubisz M, Iwanicki A, Hinc K, Kaźmierkiewicz R, Obuchowski M. 2018. Positions 299 and 302 of the GerAA subunit are important for function of the GerA spore germination receptor in *Bacillus subtilis*. *PLoS ONE*. 13(6):e0198561.

Patenty:

Patent numer Pat.240353 zgłoszony dnia 2019-06-07; Twórcy: Dąbrowska A.M., Kaźmierkiewicz R, Barabaś A.M., **Jamrozek I. B.**, Makowski M.; Zgłaszający: Uniwersytet Gdański; „3-azotowe pochodne 2,3-dideoksy-D-rybo-heksopiranozydów metylu do zastosowania jako inhibitor odwrotnej transkrypcji wirusa HIV-1”.

Konferencje:

12th International Conference on Advancements in Bioinformatics and Drug Discovery; Dublin, Irlandia; 26.-28.11.2018: “Predictions of the structure of Bacillus subtilis GerA germination receptor complex in the membrane model”, **Jamrozek I**, Przygocki P., Kaźmierkiewicz R., Obuchowski M.; poster

19th International Conference on Bacilli & Gram-Positive Bacteria; Berlin, Niemcy; 11.-15.06.2017: “Developing an in silico model of GerA spore germination receptor”; **Jamrozek I.**, Kaźmierkiewicz R., Obuchowski M.; poster

19th International Conference on Bacilli & Gram-Positive Bacteria; Berlin, Niemcy; 11.-15.06.2017: “Functional analysis of GerAA subunit of Bacillus subtilis GerA germination receptor”, Grela A, **Jamrozek I**, Iwanicki A, Hinc K, Obuchowski M; prezentacja ustna

Biotech Solutions for Health and Environment: MOBI4HEALTH Conference; Gdańsk, Polska; 12.-15.04.2016: “Development of a 3D model of GerA germination receptor in the inner spore membrane”, **Jamrozek I.**, Kaźmierkiewicz R., Obuchowski M.; poster

Załącznik nr 3

Płyta CD (znajduje się na tylnej okładce pracy) zawierająca pliki pdb modeli struktur opisanych w mojej pracy.

Lista plików zawartych na płycie CD-ROM:

- i. GerAA1.pdb — pierwszy model białka GerAA
- ii. GerAA2.pdb — drugi, ulepszony model białka GerAA
- iii. GerAB.pdb — model białka GerAB
- iv. GerAC.pdb — model białka GerAC
- v. dimerGerAA.pdb — model dimerów białka GerAA
- vi. dimerGerAB.pdb — model dimerów białka GerAB
- vii. dimerGerAC.pdb — model dimerów białka GerAC
- viii. kompleksGerAmonomery.pdb — model receptora GerA zbudowany z monomerów podjednostek
- ix. kompleksGerAdimery.pdb — model receptora GerA zbudowany z dimerów podjednostek
- x. membraneGerAmon.pdb — model receptora zbudowany z monomerów zanurzony w modelu błony fosfolipidowej
- xi. membraneGerAdim.pdb — model receptora zbudowany z dimerów zanurzony w modelu błony fosfolipidowej
- xii. GerDmonomer — model monomeru białka GerD

Projekty badawcze:

Wykonawca projektu w ramach grantu HARMONIA 6 numer 2014/14/M/NZ1/00446 pod tytułem „Mechanizm działania receptora GerA w przetrwalnikach *B.subtilis*” w latach 2015-2018.

Kierownik i wykonawca projektu pod tytułem „Przewidywanie modelu receptora kiełkowania przetrwalników GerA w błonie komórkowej” w ramach konkursu na dofinansowanie badań naukowych młodych naukowców oraz uczestników studiów doktoranckich Uniwersytetu Gdańskiego w roku 2016.

Kierownik i wykonawca projektu pod tytułem „Symulacje modeli kompleksu receptora kiełkowania GerA w modelu błony fosfolipidowej” w ramach konkursu na dofinansowanie badań naukowych młodych naukowców oraz uczestników studiów doktoranckich Uniwersytetu Gdańskiego w roku 2018.

Nagrody i wyróżnienia:

Nagroda dla najlepszej prezentacji posteru na konferencji 12th International Conference on Advancements in Bioinformatics and Drug Discovery; Dublin, Irlandia, Listopad 2018

Zespołowe osiągnięcia naukowe II stopnia wydane przez Rektora Gdańskiego Uniwersytetu Medycznego w grudniu 2019 r. za badania nad *Bacillus subtilis*.