Instytut Chemii Bioorganicznej Polskiej Akademii Nauk w Poznaniu

Magdalena Bejger

Otrzymywanie i analiza krystalograficzna wybranych białek układu roślina motylkowata – rizobium

Rozprawa doktorska
przedstawiona Radzie Naukowej
IChB PAN w Poznaniu
w celu uzyskania stopnia naukowego
doktora nauk chemicznych

Promotor: prof. dr hab. Mariusz Jaskólski

Poznań 2015

Pracę tą dedykuję moim Rodzicom,
Krystynie i Franciszkowi Bejger

Pragnę serdecznie podziękować

Profesorowi Mariuszowi Jaskólskiemu
za umożliwienie przeprowadzenia interesujących badań oraz opiekę

oraz

Profesorowi Michałowi Sikorskiemu
Profesorowi Grzegorzowi Bujaczowi
Profesorowi Tomaszowi Stępkowskiemu
Doktorowi Krzysztofowi Brzezińskiemu

Doktorowi Mirosławowi Gilskiemu
Maciejowi Kaźmierczykowi

Profesorowi Frederic Marsolais
Doktor Barbarze Imiołczyk
Damien Clavel

Doktor Agnieszce Kiliszek
Doktor Katarzynie Banaszak
Doktor Joannie Raczyńskiej
Magistrowi Tomaszowi Manszewskiemu

za pomoc i cenne dyskusje

a także:

Adrianowi
za cierpliwość i wsparcie

Badania były częściowo finansowane:

- przez MNiSW w ramach projektu „Aspekty strukturalne białek nodulacji” N N302 430534 (2008 - 2011)

- ze środków przeznaczonych na realizację zadania badawczego służącego rozwojowi młodych naukowców i uczestników studiów doktoranckich IChB PAN w Poznaniu „Badania krystalograficzne białka NodA” (2012)

Otrzymane wyniki były tematem następujących wystąpień konferencyjnych oraz publikacji naukowych:

- Bejger M., Imiołczyk B., Clavel D., Gilski M., Pająk A., Marsolais F., Jaskólski M. (2014). Na^+/K^+ exchange switches the catalytic apparatus of K-dependent plant L-asparaginase. *Acta Cryst.* D**70**, 1854-1872. Doi:10.1107/S1399004714008700

- The 15th Heart of Europe bio-Crystallography Meeting, 28-30.09.2012, Beilngries, Niemcy - komunikat ustny

- 56 Konwersatorium Krystalograficzne, 26-27.06.2014, Instytut Niskich Temperatur i Badań Strukturalnych PAN we Wrocławiu - komunikat ustny

Otrzymane wyniki zostały zdeponowane w banku struktur białkowych PDB (*ang. Protein Data Bank*) :

- kod PDB: 4pu6 - Crystal structure of potassium-dependent plant-type L-asparaginase from *Phaseolus vulgaris* in complex with K^+ cations

- kod PDB: 4pv2 - Crystal structure of potassium-dependent plant-type L-asparaginase from *Phaseolus vulgaris* in complex with K^+ and Na^+ cations

- kod PDB: 4pv3 - Crystal structure of potassium-dependent plant-type L-asparaginase from *Phaseolus vulgaris* in complex with Na^+ cations

Spis Treści

1. Wykaz stosowanych skrótów ……. 13

2. Cel pracy …….…..... 15

3. Wstęp literaturowy ……..…….. 17
3.1. Układ symbiotyczny ……..…... 17
3.1.1. Związek rośliny motylkowatej i bakterii brodawkowej ……..... 18
3.1.2. Inicjacja i rozwój symbiozy ……..... 18
3.1.3. Budowa, percepcja i znaczenie czynników Nod ……..... 20
3.2. Bakteryjne białka nodulacji ……..... 22
3.2.1. Synteza czynnika Nod ……...…….…..... 23
3.2.2. Charakterystyki wybranych białek nodulacji ……..... 25
3.2.2.1. N-acylotransferaza NodA ……..……..... 25
3.2.2.2. Deacetylaza NodB ……...…….…..... 26
3.2.2.3. Metylotransferaza NoeI ……..... 27
3.2.2.4. Karbamoilotransferazy NodU i NolO ……..…... 27
3.3. Asparaginazy ……..…... 27
3.3.1. Asparaginazy typu bakteryjnego ……..…..…… 28
3.3.2. Asparaginazy typu roślinnego ……...……..…...…… 30
3.3.2.1. Rodzina Ntn-hydrolaz ……. 30
3.3.2.2. Enzymy potasowo-zależne i potasowo-niezależne ……. 32
3.3.2.3. Aktywność L-asparaginazowa i izoaspartylopeptydazowa ……..…….. 33
3.3.3. Asparaginazy typu *Rhizobium etli* ……..…... 34

4. Materiały i metody ……..…….. 37
4.1. Systemy ekspresyjne ……. 37
4.1.1. Amplifikacja i analiza DNA ……..….. 37
4.1.2. Plazmidowe wektory ekspresyjne ……..…. 38
4.1.2.1. Wektor pMALc5x i klasyczna metoda klonowania ……..... 38
4.1.2.2. Wektor pET151/D-topo i system TOPO ……. 39
4.1.3. Metoda PIPE …….. 41
4.1.4. Ekspresja białek ……. 42
4.2. Oczyszczanie białek ……..... 43
4.3. Weryfikacja jakości preparatu białka …… 44
4.4. Krystalizacja białek ……... 45
4.5. Pomiary dyfrakcyjne ……..… 47
4.6. Strategie rozwiązywania struktur krystalograficznych ……..…... 48
4.6.1. Przetwarzanie danych dyfrakcyjnych ……..…... 49
4.6.2. Metoda podstawienia cząsteczkowego (MR) ……..….. 50
4.6.3. Metoda podstawienia izomorficznego (SIR i MIR) ……..…… 50
4.6.4. Metoda dostrojonej dyfrakcji anomalnej (SAD i MAD) ……..……. 52
4.7. Udokładnianie modelu struktury ……..……..…..….. 55

4.8. Walidacja i ocena jakości modelu ……. 57
4.8. Analiza struktur i prezentacja wyników …… 58

5. Bakteryjne białka nodulacji 59
5.1. Pięć obiektów badań w kilkunastu wersjach …….……... 59
5.2. Część doświadczalna …… 59
5.2.1. Pozyskanie materiału genetycznego rizobiów ……... 59
5.2.2. Tworzenie wektorów ekspresyjnych ……..... 60
5.2.2.1. Przygotowywanie wektorów ekspresyjnych typu pET ……... 61
5.2.2.2. Przygotowywanie wektorów ekspresyjnych typu pMAL ……... 62
5.2.2.3. Przygotowywanie wektorów ekspresyjnych typu pMB ……. 64
5.2.2.4. Weryfikacja i wybór wektorów o poprawnej konstrukcji i sekwencji …… 65
5.2.3. Pilotażowa ekspresja białek nodulacji ……... 67
5.2.3.1. Transformacja komórek gospodarza ……... 67
5.2.3.2. Hodowla komórek ……... 68
5.2.3.3. Ekspresja białek ……...…... 68
5.2.3.4. Analiza wyników ekspresji ……..... 68
5.2.4. Produkcja, oczyszczanie i krystalizacja wybranych białek nodulacji ……..... 70
5.2.4.1. NodA z *Bradyrhrizobium* sp. WM9 ze znacznikiem histydynowym …..... 71
5.2.4.2. NodA z *Sinorhrizobium meliloti* ze znacznikiem histydynowym ……... 76
5.2.4.3. NodA z *Bradyrhrizobium* sp. WM9 z białkiem fuzyjnym MBP (wersja 1.) 79
5.2.4.4. NodA z *Bradyrhrizobium* sp. WM9 z białkiem fuzyjnym MBP (wersja 2.) 84
5.2.4.5. NodB z *Bradyrhizobium* sp. WM9 z białkiem fuzyjnym MBP …….. 89
5.2.4.6. NoeI z *Bradyrhizobium japonicum* z białkiem fuzyjnym MBP …….. 92
5.3. Dyskusja wyników …… 94
5.3.1. Analiza prób ekspresji, oczyszczania i krystalizacji białka NodA ……..... 94
5.3.2. Próby ekspresji, oczyszczania i krystalizacji pozostałych białek nodulacji … 95
5.3.3. Przydatność białek nodulacji do badań krystalograficznych ……..... 96

6. Asparaginaza z *Rhizobium etli* (AnsA) …… 99
6.1. Strategia postępowania z enzymem AnsA …… 99
6.2. Część doświadczalna ……...…. 99
6.2.1. Badania wstępne …… 99
6.2.2. Produkcja białka natywnego ……..... 102
6.2.3. Produkcja białka znakowanego atomami selenu …….. 103
6.2.4. Oczyszczanie białka AnsA ……. 105
6.2.5. Krystalizacja białka AnsA ……..... 108
6.2.5.1. Poszukiwanie wstępnych warunków krystalizacji ……..... 108
6.2.5.2. Optymalizacja krystalizacji ……. 110
6.2.5.3. Próby uzyskania kryształów do badań dyfrakcyjnych ……..... 117
6.2.6. Pomiary dyfrakcji rentgenowskiej …..…….. 118
6.2.7. Weryfikacja danych dyfrakcyjnych …… 120
6.2.8. Ocena jakości preparatu białka …….…. 123
6.2.8.1. Test aktywności enzymatycznej ……..... 123

6.2.8.2. Dynamiczne rozpraszanie światła ……..……. 124
6.2.8.3. Spektrometria mas ……..…… 129
6.3. Dyskusja wyników ……...…. 129
6.3.1. Weryfikacja wydajności procesów produkcji i oczyszczania białka AnsA 129
6.3.2. Ocena jakości preparatu białka AnsA …….……... 130
6.3.3. Analiza prób krystalizacyjnych ……..... 131
6.3.4. Cechy kryształów a jakość dyfrakcji ……..... 132

7. Potasowo-zależna asparaginaza z *Phaseolus vulgaris* (PvAspG1) ……..... 135
7.1. Jeden enzym w trzech wersjach …… 135
7.2. Część doświadczalna …… 136
7.2.1. Rozwiązywanie, udokładnianie i ocena jakości struktur PvAspG1 ……... 136
7.2.1.1. Strategia, modele, użyte programy ……..... 137
7.2.1.2. Niebiałkowe składniki w strukturach kryształów ……... 139
7.2.1.3. Identyfikacja jonów metali w strukturach …….. 140
7.2.2. Analiza trzech modeli PvAspG1 …… 144
7.2.2.1. Ogólna charakterystyka struktury PvAspG1 …….. 144
7.2.2.2. Sekwencja białka w świetle gęstości elektronowej …… 148
7.2.2.3. Architektura miejsca aktywnego ……. 150
7.2.2.4. Pętla stabilizująca …… 152
7.2.2.5. Pętla aktywująca ……..... 156
7.2.2.6. Przełącznik katalityczny ……..... 159
7.3. Dyskusja wyników …… 164
7.3.1. Analiza porównawcza różnych asparaginaz typu roślinnego …… 164
7.3.1.1. Porównanie sekwencji aminokwasowych …….. 164
7.3.1.2. Porównanie struktur czwartorzędowych ……. 165
7.3.1.3. Pętle koordynujące jony metali ……... 167
7.3.1.4. Miejsce aktywne i przełącznik katalityczny ……... 170
7.3.2. Autokatalityczne dojrzewanie białka ……. 171
7.3.3. Mechanizm reakcji katalitycznej …….. 173
7.3.4. Specyficzność substratowa …… 175
7.3.5. Struktura krystaliczna PvAspG1 na tle hipotez literaturowych …… 179

8. Podsumowanie ……..…..... 183
8.1. Bakteryjne białka nodulacji …….. 183
8.2. Asparaginaza z *Rhizobium etli* (AnsA) ……. 184
8.3. Potasowo-zależna asparaginaza z *Phaseolus vulgaris* (PvAspG1) …….. 185

9. Summary ……..... 189
9.1. Bacterial nodulation proteins ……..... 189
9.2. Asparaginase from *Rhizobium etli* (AnsA) ……..…… 190
9.3. Potassium-dependent asparaginase from *Phaseolus vulgaris* (PvAspG1) …... 191

10. Bibliografia ……..…...…….... 195

1. Wykaz stosowanych skrótów

(wytłuszczoną czcionką zaznaczyłam białka badane w niniejszej pracy)

AGA	-	aspartyloglukozaminidaza
AnsA	-	**asparaginaza II z *Rhizobium etli***
At3g16150	-	asparaginaza potasowo-zależna z *Arabidopsis thaliana*
At5g08100	-	asparaginaza potasowo-niezależna z *Arabidopsis thaliana*
BVM	-	metoda weryfikacji tożsamości jonu (*ang. Bond Valence Method*)
CAPS	-	kwas N-cykloheksylo-3-aminopropanosulfonowy
CCP4	-	pakiet programów (*ang. Collaborative Computing Project Number 4*)
CMM	-	serwer weryfikujący tożsamość jonu metalu (*ang. CheckMyMetal*)
CSD	-	baza krystalograficznych struktur organicznych (*ang. Cambridge Structural Database*)
CTAB	-	surfaktant, bromek trimetylo-cetyloamoniowy
DDT	-	DL-ditiotreitol, 2,3-dihydroksybutan-1,4-ditiol
DLS	-	dynamiczne rozpraszanie światła *(ang. Dynamic Light Scattering*)
DSSP	-	program do analizy struktury drugorzędowej białek, (*ang. Definition of Secondary Structure of Proteins*)
EcAIII	-	asparaginaza typu roślinnego z *Escherichia coli*
EDTA	-	kwas etylenodiaminotetraoctowy
FPLC	-	wysokosprawna chromatografia cieczowa (*ang. Fast Performance Liquid Chromatography*)
Gly/NaOH	-	roztwór buforowy utrzymujący wysokie pH mieszanin
gpASNase3	-	asparaginaza typu roślinnego z *Cavia porcellus* (*ang. Guinea pig*)
hASNase3	-	asparaginaza typu roślinnego z *Homo sapiens (ang. Human)*
HEPES	-	kwas 4-(2-hydroksyetylo)-1-piperazynoetanosulfonowy
IPTG	-	izopropylo-β-D-tiogalaktopiranozyd
LB	-	pożywka Luria-Bretani
LjNSE1	-	asparaginaza potasowo-zależna z *Lotus japonicus*
LjNSE2	-	asparaginaza potasowo-niezależna z *Lotus japonicus*
LlA	-	asparaginaza potasowo-niezależna z *Lupinus luteus*
MBP	-	białko wiążące maltozę (*ang. Maltose Binding Protein)*
mcs	-	polilinker *(ang. multiple cloning site)*
MES	-	kwas 2-(*N*-morfolino)etanosulfonowy
MPD	-	2-metylopentano-2,4-diol
NCBI	-	centrum informacji biotechnologicznej, którego częścią jest GenBank, *(ang. National Center for Biotechnology Information)*
NDSB	-	czynnik zwiększający rozpuszczalność, kwas 3-(1-pirydyno)-1-propanosulfonowy (*ang. non detergent sulfobetaine*)

NF	-	czynnik Nod (*ang. Nod Factor*)
NodA	-	**N-acylotransferaza**
NodB	-	**deacetylaza**
NodU	-	**O-karbamoilotransferaza**
NoeI	-	**metylotransferaza**
NolO	-	**O-karbamoilotransferaza**
Ntn	-	N-końcowy nukleofil (*ang. N-terminal nucleophile*)
PBM	-	błona peribakteroidalna (*ang. peribacteroid membrane*)
PCR	-	łańcuchowa reakcja polimeryzacji nukleotydów (*ang. Polymerase Chain Reaction*)
PDB	-	bank struktur białkowych (*ang. Protein Data Bank*)
PEG	-	glikol polietylenowy
PIPE	-	metoda klonowania i mutagenezy (*ang. Polymerase Incomplete Primer Extension*)
PMF	-	peptydowy odcisk palca, technika analityczna, (*ang. Peptide Mass Fingerprinting*),
PMSF	-	fluorek α-toluenosulfonylu
PvAspG1	-	**asparaginaza potasowo-zależna z *Phaseolus vulgaris***
PvAspGT2	-	asparaginaza potasowo-niezależna z *Phaseolus vulgaris*
p.z.	-	par zasad, jednostka określająca długość dwuniciowego łańcucha DNA
r.m.s.d.	-	średnie kwadratowe odchylenie (*ang. root-mean-square deviation*)
rpm	-	obroty na minutę *(ang. revolutions per minute*)
SAD	-	dyfrakcja anomalna przy jednej długości fali (*ang. Single-wavelength Anomalus Diffraction*)
SDS	-	dodecylosiarczan sodu
SDS-PAGE	-	elektroforeza w żelu poliakrylamidowym w obecności SDS
Se-Met	-	selenometionina
SOC	-	pożywka zawierająca: ekstrakt drożdżowy, trypton, glukozę, NaCl, KCl, $MgCl_2$, $MgSO_4$
SSM	-	narzędzie z programu COOT porównujące różne modele białek
Tas1	-	taspaza
TB	-	pożywka Terrific Broth
TCA	-	kwas trichlorooctowy
TCEP	-	tris(2-karboksyetylo)fosfina
TEV^{His}	-	proteaza TEV zawierająca znacznik histydynowy
TLS	-	model drgań anizotropowych bryły sztywnej (*ang. translation, libration, screw motions*)
TMAO	-	N-tlenek trimetyloaminy
Tris	-	2-amino-2-(hydroksymetylo)-1,3-propanodiol
Triton X-100	-	eter izooktylofenylowy polioksoetylenu (10)
YMB	-	pożywka Yeast Mannitol Broth

2. Cel pracy

„Otrzymywanie i analiza krystalograficzna wybranych białek układu roślina motylkowata – rizobium" jest praca doktorską, która tematycznie i metodycznie łączy trzy zadania składowe, które zostały wyróżnione w układzie rozprawy. Każde z zadań stanowi fragment oddzielnego projektu.

Dalekosiężnym celem wszystkich trzech projektów jest/było poznanie struktur białek z układu symbiotycznego roślina - bakteria. Wszystkie trzy projekty łączy także zastosowana metoda badawcza, czyli rentgenowska analiza strukturalna. Cały proces określania struktury przestrzennej białek składa się jednak z wielu etapów obejmujących: projektowanie i tworzenie systemów ekspresyjnych, produkcję białek w komórkach gospodarza, oczyszczanie białek metodami chromatografii, krystalizację, pomiary dyfrakcji rentgenowskiej, rozwiązywanie, udokładnianie oraz interpretację struktur. Gdy włączałam się do kolejnych projektów, postęp prac nad nimi był w różnym stopniu zaawansowany, stąd zadania jakie otrzymałam dotyczą różnych etapów całego procesu określania struktur.

W pierwszym z projektów, który można określić jako „genomika strukturalna biosyntezy czynnika Nod", obiektem badań była grupa kilkunastu tzw. białek nodulacji, pochodzących z symbiotycznych bakterii brodawkowych. Do tej pory zidentyfikowano około sześćdziesiąt różnych bakteryjnych genów nodulacyjnych, natomiast struktury białek z tej grupy znane są tylko trzy: NodZ (**Brzeziński *et al.*, 2007**), NodS (**Cakici *et al.*, 2010**) i NolR (**Lee *et al.*, 2014**). Moim zadaniem w tym projekcie było przetestowanie pięciu wybranych białek nodulacji pod kątem ich przydatności do badań krystalograficznych, które w dalszej perspektywie pozwoliłyby na poszerzenie wiedzy strukturalnej na temat układów symbiotycznych roślina motylkowata - bakterie rizobium. Zajmowałam się białkami nodulacji: NodA, NodB, NodU, NolO i NoeI. Szczegółowy zakres prac obejmował (1) uzyskanie plazmidowych wektorów ekspresyjnych, (2) produkcję białek w komórkach *E. coli*, (3) oczyszczanie i (4) krystalizację białek.

W projektach drugim i trzecim, do których dołączyłam równocześnie, obiektami badań były/są asparaginazy pochodzące z organizmów symbiotycznych, konkretnie z pary symbiontów *Phaseolus vulgaris* i *Rhizobium etli*.

Asparaginaza z *Rhizobium etli* (białko nazwane AnsA) jest przedstawicielem słabo poznanej, nie posiadającej zdefiniowanego modelu struktury, rodziny enzymów typu *R. etli*, którą cechuje znaczna odmienność sekwencyjna w porównaniu do innych asparaginaz. Moim zadaniem było umożliwienie podjęcia próby rozwiązania pierwszej struktury tego typu enzymu, a konkretnie, przeprowadzenie obiektu badań przez te wszystkie etapy projektu, które dotyczą pracy z preparatem białkowym. Prace rozpoczęłam w oparciu o przygotowany wcześniej w naszej pracowni wektor ekspresyjny (**Kaźmierczyk, praca magisterska**) i miałam zakończyć rejestrując zbiory danych pozwalających wykorzystać różne strategie rozwiązywania struktury. Szczegółowy zakres moich prac obejmował (1) produkcję rozpuszczalnych białek, (2) opracowanie protokołu oczyszczania, (3) znalezienie warunków krystalizacyjnych oraz (4) uzyskanie danych dyfrakcyjnych. Projekt dotyczący asparaginazy z *Rhizobium etli* jeszcze się nie zakończył i obecnie uczestniczą w nim kolejne osoby pracujące nad rozwiązaniem struktury tego wyjątkowego enzymu.

Pracę nad drugą asparaginazą - potasowo-zależnym enzymem typu roślinnego z *Phaseolus vulgaris* (PvAspG1) - rozpoczęłam z kolei od etapu, na którym zakończyłam zadanie dotyczące AnsA, wychodząc z zarejestrowanych w naszej pracowni danych dyfrakcyjnych. Zadaniem moim w tym projekcie było (1) rozwiązać, (2) udokładnić, a następnie (3) szczegółowo przeanalizować strukturę białka. Celem zadania było wyjaśnienie podstaw molekularnych zależności aktywności enzymatycznej asparaginazy z *Phaseolus vulgaris* od jonów potasu.

3. Wstęp literaturowy

3.1. Układ symbiotyczny

Prawidłowe funkcjonowanie i rozwój organizmów żywych jest uzależnione od dostępności przyswajalnych form azotu. Rośliny przystosowane są do pobierania azotu z gleby w postaci jonów azotanowych (NO_3^-) i amonowych (NH_4^+). W warunkach niedoboru azotu w podłożu niektóre rośliny mogą wejść w symbiozę z mikroorganizmami zdolnymi do wiązania, niedostępnego dla roślin azotu atmosferycznego (**Santi *et al.*, 2013**). W ten sposób pozyskują nowe źródło brakującego makroelementu, w zamian oferując produkty fotosyntezy, które dostarczają bakteryjnemu symbiontowi. Taka symbiotyczna współpraca jest możliwa pomiędzy roślinami motylkowatymi / bobowatymi (rodzina *Fabaceae*), a bakteriami glebowymi z rodzajów *Rhizobium, Bradyrhizobium, Sinorhizobium, Mesorhizobium, Azorhizobium* i *Allorhizobium* (potocznie nazywanymi rizobiami).

Ustanowienie symbiozy pomiędzy roślinami motylkowatymi, a bakteriami glebowymi jest jednoznaczne z wytworzeniem nowego organu roślinnego – brodawki korzeniowej, która zostaje skolonizowana przez bakterie. Wewnątrz roślinnego gospodarza rizobia (po przekształceniu w bakteroidy) asymilują azot atmosferyczny dzięki aktywności bakteryjnego kompleksu nitrogenazy. Rozwój funkcjonalnej brodawki korzeniowej wymaga jednak aktywacji i regulacji licznych genów symbiotycznych, zarówno lokalnie jak i w skali całej rośliny oraz wiąże się z dostawami ogromnych ilości energii i materiałów budulcowych. Nawiązanie symbiozy rozpoczyna się od wymiany między symbiontami precyzyjnych sygnałów molekularnych, które umożliwiają odpowiednie dobranie i wzajemne rozpoznanie partnerów. Najważniejszym związkiem uczestniczącym w tym porozumiewawczym dialogu symbiontów jest produkowany przez bakterie czynnik Nod (*ang. Nod Factor*, NF). Zarówno inicjacja jak i rozwój symbiozy jest procesem bardzo złożonym. Wymaga odpowiedniej sygnalizacji, synchronizacji licznych procesów biochemicznych oraz ekspresji informacji genetycznej obu symbiontów w odpowiednim czasie i miejscu.

3.1.1. Związek rośliny motylkowatej i bakterii brodawkowej

W układ symbiotyczny z rizobiami może wejść ok 12 000 różnych gatunków roślin motylkowatych (**Maroti i Kondorosi, 2014**), przy czym oddziaływania pomiędzy symbiontami są w różnym stopniu specyficzne. Większość rizobiów wchodzi w symbiozę tylko z nielicznymi roślinami, choć niektóre gatunki bakterii wykazują bardzo szeroki zakres gospodarzy, jak na przykład szczep *Sinorhizobium* sp. NGR234, który tworzy brodawki na ok 112 gatunkach roślin motylkowatych (**Pueppke i Broughton, 1999; Iantcheva *et al.*, 2013**). Podobnie różne gatunki roślin mogą być gospodarzem dla zróżnicowanego, aczkolwiek precyzyjnie dobranego wachlarza bakteryjnych symbiontów. Fasola zwyczajna *(Phaseolus vulgaris*) dla przykładu może być gospodarzem dla szeregu rizobiów (*Rhizobium leguminosarum, Rhizobium etli, Rhizobium gallicum, Rhizobium tropici, Rhizobium giardinii*) (**Sessitsch *et al.*, 1997**).

Dzięki symbiozie roślin motylkowatych z bakteriami brodawkowymi możliwa jest asymilacja azotu atmosferycznego w całkowitej ilości ok. 200 mln ton N_2 rocznie (**Graham i Vance, 2003**). Wiele gatunków roślin motylkowatych ma ogromne i niepodważalne znaczenie gospodarcze. Rośliny te uprawiane są na ok. 180 mln ha, a plony stanowią 27% światowej produkcji (**Graham i Vance, 2003**). Dostarczają ok. 35% przetwarzanego na świecie oleju roślinnego, są żywnością dla człowieka i paszą dla zwierząt. Szacuje się, że co trzeci atom azotu w ludzkim organizmie został wychwycony z powietrza przez symbiozę (**Gresshoff *et al.*, 2014**). Dla rolnictwa uprawa roślin takich jak soja, orzeszek ziemny, fasola, lucerna, koniczyna, łubin czy groch oznacza również mniejsze zużycie nawozów sztucznych oraz poprawę fizycznych i chemicznych parametrów gleb. Rośliny motylkowate wykorzystywane są w przemyśle spożywczym i chemicznym, medycynie, leśnictwie i jako rośliny ozdobne.

3.1.2. Inicjacja i rozwój symbiozy

Skomplikowany proces ustanawiania symbiozy pomiędzy roślinami motylkowatymi i bakteriami glebowymi został szczegółowo opisany w wielu znakomitych publikacjach przeglądowych (**Hamel i Beaudoin, 2010; Kouchi *et al.*, 2010; Murray, 2011; Popp i Ott, 2011; Oldroyd, 2013**). Proces ten rozpoczyna się wymianą pomiędzy symbiontami sygnałów o charakterze chemicznym w celu zidentyfikowania odpowiedniego partnera. Roślinnymi związkami sygnałowymi są

flawonoidy, które wydzielone do gleby oddziałują na bakterie, żyjące w ścisłym związku z korzeniami roślin w strefie zwanej ryzosferą. Flawonoidy powodują aktywację bakteryjnych genów symbiotycznych (**Moscatiello *et al.*, 2010; Abdel-Lateif *et al.*, 2012**), dzięki którym rizobia rozpoczynają produkcję własnego związku sygnałowego - czynnika Nod (Rozdział 3.2). Wydzielony poza komórkę bakteryjną NF jest z kolei specyficznie rozpoznawany przez roślinę i powoduje aktywację wielu różnorodnych mechanizmów prowadzących do ustanowienia symbiozy (Rozdział 3.1.3) Chemiczny dialog pomiędzy symbiontami angażuje także inne cząsteczki sygnałowe (np. białka wydzielane poza komórki, polisacharydy występujące na powierzchni bakteryjnych symbiontów) i jest kontynuowany w trakcie całego procesu rozwoju symbiozy (**Fraysse *et al.*, 2003; Cooper, 2007; Downie, 2010**).

Rozpoznanie bakteryjnych symbiontów aktywuje równocześnie dwa roślinne programy: jeden prowadzący do rozwoju brodawki korzeniowej, drugi umożliwiający zainfekowanie nowego organu bakteriami symbiotycznymi. W procesie formowania się brodawki pierwszym etapem jest wykształcenie się zawiązka. U roślin strefy umiarkowanej zawiązek tworzą dzielące się komórki wewnętrznej warstwy kory pierwotnej korzenia doprowadzając do powstania brodawki typu niezdeterminowanego. Z kolei u roślin tropikalnych zawiązki zlokalizowane są w zewnętrznej warstwie kory i przekształcają się w brodawki zdeterminowane (**Kondorosi *et al.* 2013**).

Wniknięcie bakterii do organizmu roślinnego rozpoczyna się, gdy młode włośniki skręcając się obrastają przyczepione do ich powierzchni rizobia. W miejscu styku powstaje wgłębienie roślinnej błony komórkowej i tworzy się nić infekcyjna, której zadaniem jest przetransportowanie zamkniętych w jej wnętrzu rizobiów do tworzącego się wewnątrz korzenia zawiązka brodawki. Znacznie rzadziej następuje infekcja poprzez pęknięcia i przerwy w epidermie korzenia (**Maroti i Kondorosi, 2014**). Nić infekcyjna uwalnia bakterie do cytoplazmy komórek zawiązka, gdzie zostają otoczone przez błonę cytoplazmatyczną rośliny, nazywaną błoną peribakteroidalną (*ang. peribacteroid membranę,* PBM), tworząc wewnątrz komórki gospodarza tzw. symbiosomy. Zadaniem błony PBM jest ochrona bakterii przed reakcjami obronnymi rośliny, umożliwienie wymiany cząsteczek sygnałowych oraz przepływ metabolitów (np. związków węgla i azotu).

W dojrzałych brodawkach symbiosomy zawierają rizobia przekształcone w bakteroidy, które przystosowane są do endosymbiotycznego trybu życia. Obecny w bakteroidach kompleks nitrogenazy katalizuje redukcję N_2 do NH_3. Wyprodukowany

amoniak jest następnie eksportowany z symbiosomu do cytoplazmy roślinnej, gdzie jest przekształcany przez enzymy gospodarza do takiej formy, w której może być transportowany poprzez wiązki przewodzące, z brodawki do nadziemnych organów roślinnych (**Udvardi i Poole, 2013**).

W roślinach głównym związkiem transportującym i magazynującym azot jest asparagina (**Azevedo *et al.*, 2006**). Za jej produkcję odpowiada syntetaza asparaginy (*ang. asparagine synthetase*), natomiast asparaginaza (*ang. asparaginase*) przeprowadza katabolizm uwalniając amoniak (Rozdział 3.3), który jest re-asymilowany poprzez rozmaite szlaki metaboliczne.

3.1.3. Budowa, percepcja i znaczenie czynników Nod

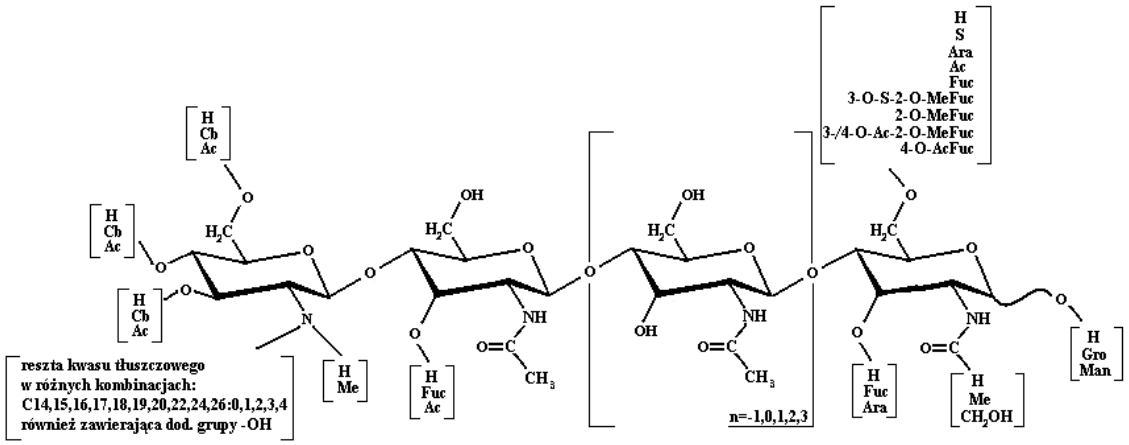
Pomijając nieliczne wyjątki (**Giraud *et al.*, 2007**) czynniki Nod są niezbędne dla ustanowienia symbiotycznego związku. Umożliwiają wzajemne rozpoznawanie się partnerów, co daje pewność efektywnej symbiozy i wyklucza możliwość wniknięcia do rośliny bakteryjnych intruzów obficie występujących w glebie. Co więcej, wykazano, że wyizolowane i oczyszczone czynniki Nod są zdolne bez udziału bakterii rozpocząć proces symbiozy (m.in. wywołują deformacje włośników inicjując infekcję, aktywują roślinne geny zaangażowane w rozwój brodawki, indukują podziały komórkowe prowadzące do formowania się zawiązka brodawki) (**D'Haeze i Holsters, 2002**). Czynniki Nod produkowane są również przez bakterie podróżujące już wśród tkanek roślinnych do zawiązka brodawki, ponieważ są one niezbędne dla utrzymania wzrostu nici infekcyjnej (**Downie, 2010**).

Rizobia potrafią syntetyzować czynniki Nod o wielu różnorodnych modyfikacjach. Jeden rodzaj bakterii symbiotycznej może syntetyzować od 2 do około 60 rodzajów tej cząsteczki, w różnych proporcjach ilościowych. Dzięki tej różnorodności NF decydują o specyficzności zakażania konkretnych odmian roślin motylkowatych.

Ogólny schemat budowy chemicznej wszystkich poznanych dotąd czynników Nod jest podobny (Rysunek 3.1) (**D'Haeze i Holsters, 2002**). Są one lipochitooligo-sacharydami. Szkielet główny czynnika składa się z dwóch do sześciu jednostek N-acetylo-D-glukozoaminy połączonych wiązaniami β-1,4-glikozydowymi. Koniec nieredukujący oligosacharydu chitynowego jest N-acylowany w pozycji C-2 resztą kwasu tłuszczowego, zróżnicowaną pod względem długości łańcucha węglowego

(najczęściej C_{16}, C_{18}, C_{20}) i liczby wiązań nienasyconych (od zera do czterech). Na końcu nieredukującym często występują dodatkowo grupy N-metylowe, O-acetylowe lub/i O-karbamoilowe. Na przeciwnym, redukującym końcu cząsteczki pojawiają się najczęściej takie podstawniki jak grupa acetylowa, sulfonowa, lub/i takie reszty cukrowe jak L-fukoza. Szkielet cukrowy może zawierać również inne rzadziej występujące podstawniki. Wszystkie możliwe podstawniki oraz ich rozmieszczenie w strukturze cząsteczki Nod przedstawia Rysunek 3.1.

Rysunek 3.1. Ogólny schemat struktury czynnika Nod z uwzględnieniem modyfikacji (w nawiasach); użyte skróty: S- reszta sulfonowa, Cb- reszta karbamoilowa, Ac- reszta acetylowa, Me- reszta metylowa, Fuc- reszta fukozylowa, Ara- reszta arabinozylowa, Gro- reszta glicerolowa, Man- reszta mannozylowa (**D'Haeze i Holsters, 2002**).



Czynniki Nod rozpoznawane są przez roślinę za pomocą dwóch transmembranowych receptorów LysM-RLK (*ang. LysM receptor like kinases*), które znajdują się w epidermie korzenia. Białka te posiadają zewnątrzkomórkowe domeny typu LysM i wewnątrzkomórkowe domeny kinazowe (**Gough i Jacquet, 2013**), prawdopodobnie funkcjonują jako hetero-kompleks i razem odpowiadają za przekazywanie sygnału NF (**Oldroyd, 2013**). Receptory te to LYK3 i NFP zidentyfikowane w *Medicago truncatula* (**Amor *et al.*, 2003; Arrighi *et al.*, 2006; Limpens *et al.*, 2003; Smit *et al.*, 2007**), NFR1 i NFR5 w *Lotus japonicus* (**Madsen *et al.*, 2003; Radutoiu *et al.*, 2003)** lub ich odpowiedniki w innych gatunkach roślin (**Zhukov *et al.*, 2008; Indrasumunar *et al.*, 2010, 2011**). Co ciekawe, tego typu receptory służą roślinie prawdopodobnie również do rozpoznawania symbiontów mikorytycznych oraz patogenów (**Oldroyd, 2013**).

Percepcja czynników Nod wywołuje w komórkach roślinnych oscylacyjne zmiany w poziomie stężenia jonów wapnia (*ang. calcium spiking*), po czym następuje ekspresja roślinnych białek symbiotycznych. Szlak transdukcji sygnału NF nie jest w pełni

poznany, zidentyfikowano jednak szereg protein zaangażowanych w ten proces. Są to między innymi homologiczne białka SYMRK (*ang. symbiosis receptor kinase)* z *L. japonicus* i NORK z *M. truncatula* (**Endre *et al.*, 2002; Stracke *et al.*, 2002**) oraz białko CCaMK (*ang. calcium- and calmodulin- dependent serine/treonine protein kinase*) (**Levy *et al.*, 2004**) i białka NSP1 i NSP2 (*ang. nodulation signaling pathways*) tworzące hetero-kompleks czynników transkrypcyjnych i łączące się z promotorami genów indukowanych czynnikiem Nod (**Smit *et al.*, 2005; Kalo *et al.*, 2005**).

3.2. Bakteryjne białka nodulacji

Rizobia mają względnie duże genomy co prawdopodobnie związane jest z ich zdolnością do życia w dwóch zupełnie różnych niszach, zarówno w narażeniu na zróżnicowane czynniki stresowe środowiska gleby, jak i w symbiozie z roślinami motylkowatymi (**Downie, 2010**). Genom *Sinorhizobium meliloti*, dla przykładu, jest złożony z chromosomu (3.65 Mbp), dwóch mega-plazmidów pSymA i pSymB (1.35 i 1.68 Mbp), które zawierają większość genów symbiotycznych, oraz dodatkowo średniej wielkości plazmidów pomocniczych, stąd może on zawierać nawet do 9 000 genów **(Maroti i Kondorosi, 2014**).

Bakteryjne geny uczestniczące w symbiozie posiadają zróżnicowane funkcje, odpowiadają między innymi za proces wiązania azotu atmosferycznego (*fix, nif*), czy syntezę powierzchniowych egzopolisacharydów (*exo, ndv*) (**Spaink *et al.*, 1998**). Grupa genów *nod* zawiera najważniejsze dla symbiozy geny nodulacji (*ang. nodulation genes*). Do tej pory odkryto i zidentyfikowano ponad pięćdziesiąt różnych genów nodulacji, którym przypisywano czteroliterowe nazwy składające się z członu *nod, nol* lub *noe* oraz kolejnych liter alfabetu. Grupę tę tworzą głównie geny uczestniczące w biosyntezie czynnika Nod. Trzy z nich oznaczone jako *nodA, nodB* i *nodC* występują u wszystkich rizobiów, skutkiem czego określane są jako geny *com-nod*, tj. wspólne geny *nod* (*ang. common nod genes*), natomiast geny *hsn-nod* (*ang. host specificity of nodulation*), np. *nodU, nodS, nodZ, nolO*, obecne są w odmiennych aranżacjach u różnych gatunków bakterii i powodują powstawanie różnorakich wersji czynnika Nod (**Perret *et al.*, 2000**). Pozostałe geny z tej grupy odpowiadają za sekrecję NF (*nodI, nodJ*), kodują białka regulacyjne (np. *nodD, nodV, nodW, nolR*) lub posiadają inne zdefiniowane lub nieokreślone funkcje.

3.2.1. Synteza czynnika Nod

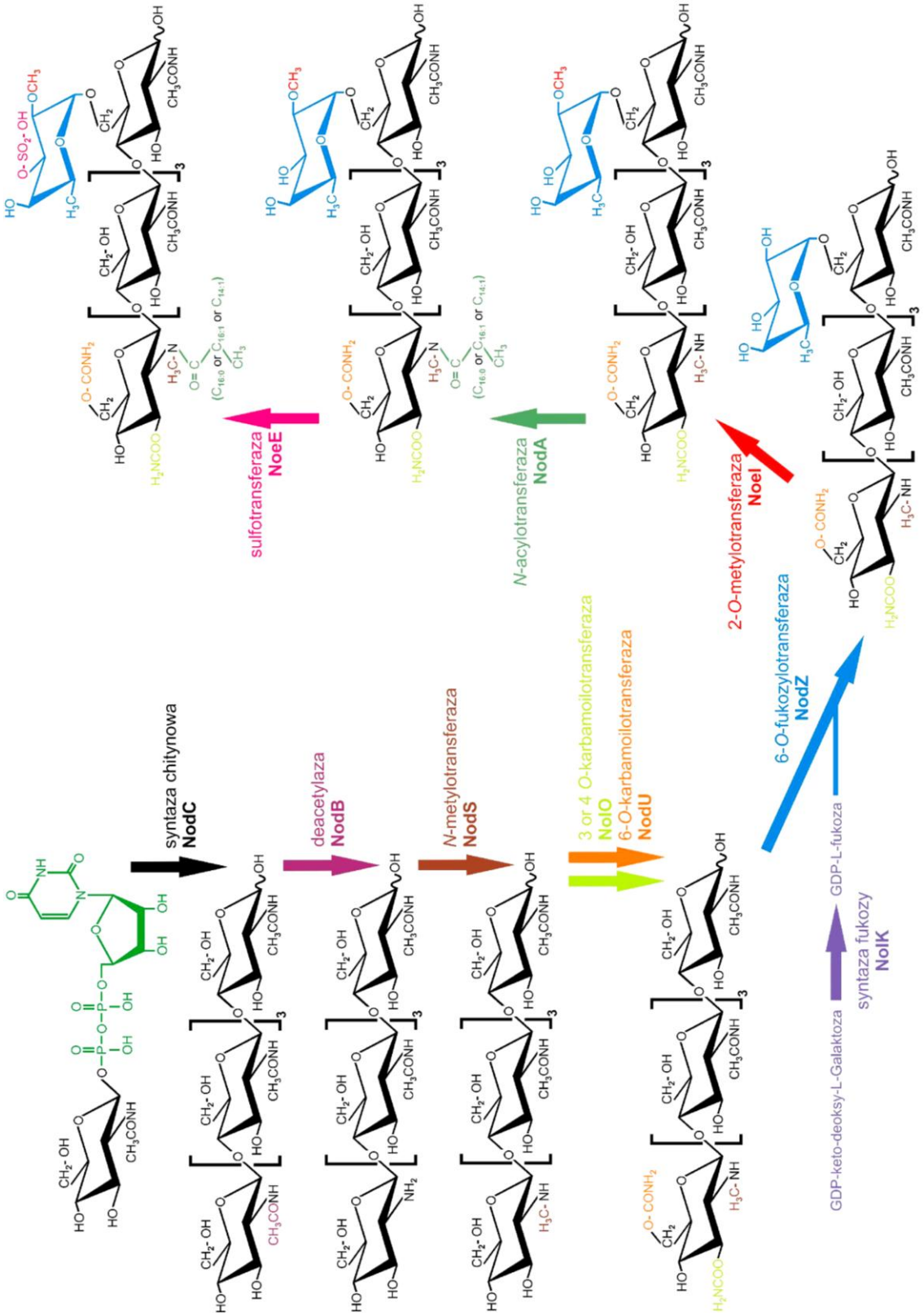
Ekspresja bakteryjnych genów *nod* indukowana jest roślinnymi cząsteczkami sygnałowymi – flawonoidami. Powodują one aktywację bakteryjnego białka NodD należącego do rodziny regulatorów transkrypcyjnych typu LysR. Białko NodD wiąże się do wysoce zachowawczych motywów DNA (*ang. nod boxes*) występujących w regionach promotorowych obszarów nodulacyjnych (**Schlaman *et al.*, 1992**). Produkty białkowe genów nodulacyjnych uczestniczą w syntezie bakteryjnego związku sygnałowego - czynnika Nod.

Synteza czynnika Nod polega na następujących po sobie reakcjach enzymatycznych. W różnych organizmach schemat powstawania jest oczywiście nieco odmienny i zależny od białek, jakimi dysponuje dany organizm. Trzy podstawowe enzymy u wszystkich rizobiów działają zawsze w tej samej kolejności (choć niekoniecznie bezpośrednio po sobie): NodC > NodB > NodA (Rysunek 3.2). Pozostałe enzymy, odpowiadające za ściśle określone modyfikacje czynnika oraz produkcję substratów do tych modyfikacji, aktywne są na różnych etapach syntezy. Procesem rządzi prawdopodobnie sam czynnik Nod – jako substrat i produkt kolejnych reakcji.

Szkielet oligosacharydowy cząsteczki Nod powstaje z D-acetyloglukozamino-1-UDP. Odpowiedzialna za ten proces jest N-acetyloglukozaminotransferaza (syntaza chitynowa) NodC, która prawdopodobnie jest białkiem błonowym. NodC wydłuża chitynowy łańcuch od końca nieredukującego determinując długość szkieletu czynnika Nod (**Kamst *et al.*, 1997**). Deacetylaza NodB usuwa resztę N-acetylową z nieredukującego końca cząsteczki (**John *et al.*, 1993; Spaink *et al.*, 1994**) przygotowując łańcuch chitynowy do kolejnych modyfikacji. N-acylotransferaza NodA przyłącza resztę kwasu tłuszczowego (acylową) do nieredukującego końca C-2 cząsteczki (**Rohrig *et al.*, 1994)**. Do szkieletu chitynowego przyłączane są dwa podstawowe typy kwasów tłuszczowych: powszechnie występujące w czynniku Nod nasycone lub jednonienasycone reszty kwasów tłuszczowych oraz rzadziej, związki nienasycone (zawierające od 2 do 4 wiązań podwójnych), które znaleziono w szczepach zawierających białka NodE i NodF (**Denarie *et al.*, 1992**).

O-acetylotransferaza NodL katalizuje z kolei acetylację nieredukującej reszty szkieletu chinowego w pozycji C-6 (**Ardourel *et al.*, 1994**). Grupy karbamoilowe są przyłączane do nieredukujacej reszty cukrowej przy udziale O-karbamoilotransferaz NodU (amidacja zachodzi w pozycji C-6 glukozaminy) i/lub NolO (amidacja w pozycji

Rysunek 3.2. Fragment szlaku biosyntezy czynnika Nod w *Rhizobium* sp. NGR234. Schemat przygotowałam na podstawie (**Broughton *et al.*, 2000**).



C-3 lub C-4 glukozaminy) (**D'Haeze *et al.*, 1999**; **Jabbouri *et al.*, 1998**). N-metylotransferaza NodS zależna od S-adenzylo-L-metioniny przeprowadza reakcję N-metylacji szkieletu oligosacharydowego po deacetylacji nieredukujacego końca chityny przez NodB (**Geelen *et al.*, 1995**). Fukozylotransferaza NodZ przeprowadza fukozylację w pozycji C-6 końca redukującego szkieletu chitynowego jeszcze przed przyłączeniem kwasu tłuszczowego (**Quesada-Vincens *et al.*, 1997**). Taka modyfikacja spotykana jest przeważnie u rizobiow o szerokim zakresie gospodarza. Reszta fukozylowa może być dodatkowo metylowana w pozycji C-2, sulfonowana przy C-3 lub acetylowana na weglu C-4. Wszystkie te modyfikacje dokonywane są przez specyficzne dla fukozy transferazy: metylową (NoeI), sulfonową (NoeE) (**Quesada-Vincens *et al.*, 1998**), lub acetylową (NolL). NodX przeprowadza acetylację w pozycji C-6 końca redukującego. Obecność grup sulfonowych związana jest z aktywnością sulfotransferazy NodH, która dokonuje sulfonowania w pozycji C-6 końca redukującego cząsteczki. Donorem grupy sulfonowej dla tej reakcji jest 3'-fosfoadenozyno-5'-fosfosulfonian, który syntetyzowany jest przez białka NodP i NodQ.

3.2.2. Charakterystyki wybranych białek nodulacji

3.2.2.1. N-acylotransferaza NodA

W przypadku białka NodA (ok. 24 kDa) stwierdzono, że występuje ono w cytozolu (**Schmidt *et al.*, 1986**), jak również we frakcjach zawierających błony komórkowe (**Egelhoff i Long, 1985**). Prawdopodobnie nie jest białkiem trans-membranowym, gdyż jego sekwencja jest relatywnie hydrofilowa. Możliwe, że powierzchnia NodA zawiera pewne obszary hydrofobowe, którymi białko mogłoby „przyklejać się" do błon. Struktura NodA nie jest znana, a funkcja została określona za pomocą badań biochemicznych (**Atkinson *et al.*, 1994)**. Białko jest N-acylotransferazą, której zadaniem jest „udekorowanie" węgla C-2 końca nieredukującego czynnika Nod **(Rohrig *et al.*, 1994**). Dwa podstawowe typy kwasów tłuszczowych dodawane są przez NodA do cząsteczki oligosacharydu chitynowego: (I) nasycone lub jednonienasycone reszty kwasów tłuszczowych oraz (II) związki nienasycone (zawierające od 2 do 4 wiązań podwójnych) (**Denarie *et al.*, 1992**).

Proponuje się, że donorami substratów acylowych dla NodA są białka ACP (*ang. acyl carrier proteins*) (**Ritsema *et al.*, 1997**). Rizobia produkujące czynniki Nod

zawierające wielonienasycone kwasy tłuszczowe (np. *R. leguminosarum, S. meliloti)* posiadają geny *nodE i nodF* (**Spaink *et al.*, 1998**). Białko NodF jest homologiczne z białkami ACP, a NodE z syntazami kwasów tłuszczowych (*ang. β-keto acyl synthases* / *β-aceto acetyl synthase*) (**Ritsema *et al.*, 1998; Shearman *et al.* 1986; Bloemberg *et al.*, 1995**). Zatem w szczepach tych NodA najprawdopodobniej współpracuje z acylo-NodF jako donorem substratu. W rizobiach produkujących czynniki Nod udekorowane nasyconymi grupami acylowymi (np. *Bradyrhizobium*) białek NodE i NodF nie odnaleziono i w tych przypadkach donorami substratów dla NodA są prawdopodobnie enzymy metabolizmu podstawowego (*ang. housekeeping enzymes*).

Różne białka NodA (z różnych gatunków rizobiów) są specyficzne wobec konkretnych kwasów tłuszczowych stanowiących ich substraty. Przykładowo, zamiana genu *nodA* występującego w *S. meliloti* na gen pochodzący z *R. tropici* powoduje, że produkowane są czynniki Nod acylowane kwasem wakcenowym (*ang. vaccenic acid*) zamiast charakterystycznymi dla *S. meliloti* kwasami nienasyconymi C16 (*ang. C16 unsaturated*) lub hydroksylowanymi (*ang. hydroxylated fatty acid*) (**Debelle *et al.* 1996**). Podobnie białko NodA z *Bradyrhizobium* sp. ANU289 ulegające ekspresji w *R. leguminosarum* biovar *viciae* nadal uczestniczy w produkcji czynników Nod, jednak nie może przenieść wielonienasyconych kwasów tłuszczowych na cząsteczkę chitynowego oligosacharydu (**Ritsema *et al.*, 1996**). Może to być spowodowane tym, że ten konkretny enzym NodA nie jest zdolny do rozpoznania (I) wielonienasyconych kwasów tłuszczowych lub/i (II) wyspecjalizowanego białka transportowego (ACP) NodF, które obecne jest w *R. leguminosarum* biovar *viciae*. Co ciekawe, transfer genów *nodE* i/lub *nodF* z jednego szczepu do innego, również ma wpływ na typ podstawianej grupy acylowej (**Demond *et al.*, 1993**). Tak więc **Ritsema *et al.* (1996**) proponują, że obecność wielonienasyconych kwasów tłuszczowych w czynnikach Nod jest determinowana na dwóch poziomach, obejmujących udział transferazy NodA i białek pomocniczych NodE i NodF. Badania pokazują jednak, że również inne białka mogą być na zaangażowane w acylację czynników Nod (**Schlaman *et al.*, 2006**).

3.2.2.2. Deacetylaza NodB

NodB jest białkiem o masie ok. 24 kDa, zlokalizowanym w cytozolu *S. meliloti* (**Schmidt *et al.*, 1988**). Jego sekwencja jest homologiczna z sekwencją deacetylazy chitynowej występującej u grzyba *Mucor rouxii* (**Kafetzopoulos *et al.*, 1993**). Funkcja

NodB została potwierdzona również za pomocą badań biochemicznych (**John *et al.*, 1993; Spaink *et al.*, 1994**). Białko NodB potrafi usunąć grupę N-acetylową z oligosacharydu posiadającego od dwóch do sześciu podjednostek, przy czym reakcja następuje tylko na jednej, końcowej reszcie chitynowego łańcucha. NodB jest nieaktywny wobec monomerycznej cząsteczki N-acetylo-D-glukozaminy (**John *et al.*, 1993**).

3.2.2.3. Metylotransferaza NoeI

Białko NoeI (ok. 24 kDa) jest odpowiedzialne za 2-O-metylację przyłączonej do szkieletu czynnika Nod reszty fukozylowej (**Jabbouri *et al.*, 1998**). Potwierdzono obecność tego białka w szczepach *Rhizobium* sp. NGR234 (**Freiberg *et al.*, 1997**) oraz *Bradyrhizobium japonicum* (**Gottfert *et al.*, 2001**).

3.2.2.4. Karbamoilotransferazy NodU i NolO

Koniec nieredukujący czynnika Nod dekorowany jest grupami karbamoilowymi. O-karbamoilotransferaza NodU (ok. 60 kDa) przeprowadza amidację w pozycji C-6 glukozaminy, natomiast O-karbamoilotransferaza NolO (ok. 75 kDa) w pozycji C-3 lub C-4 (**Jabbouri *et al.*, 1995, 1998; D'Haeze *et al.*, 1999; Treilhou *et al.*, 2000**).

3.3. Asparaginazy

Asparaginazy (EC 3.5.1.1) są amidohydrolazami, które katalizują hydrolizę wiązania amidowego łańcucha bocznego asparaginy wytwarzając asparaginian i amoniak. Enzymy te są obecne prawie we wszystkich organizmach żywych, zarówno w komórkach eukariotycznych jak i prokariotycznych. Nie jest to jednak grupa jednolita pod względem strukturalnym i funkcjonalnym. Analiza sekwencji aminokwasowej enzymów posiadających aktywność asparaginazową pozwoliła wyróżnić trzy grupy tych białek: asparaginazy typu bakteryjnego, asparaginazy typu roślinnego i asparaginazy typu *Rhizobium etli* (**Borek i Jaskólski, 2001**). Tradycyjne nazewnictwo tych grup nie pokrywa się jednak z występowaniem białek danego rodzaju, np. organizm ludzki posiada sekwencje asparaginaz zaliczane do typu bakteryjnego jak i roślinnego. Okazuje się także, że czasem enzymy zdefiniowane pierwotnie jako

asparaginazy posiadają inną dominującą aktywność *in vitro* i powinny być sklasyfikowane inaczej (**Borek *et al*. 2004**). Mechanizm działania asparaginaz jest w każdej z grup nieco odmienny, jednak zawsze podobny do mechanizmu proteaz serynowych (**Dodson i Wlodawer, 1998**).

3.3.1. Asparaginazy typu bakteryjnego

Badania nad enzymami pochodzącymi z *Escherichia coli* pozwoliły wyróżnić dwie klasy asparaginaz typu bakteryjnego (**Campbell *et al*., 1967**). Oba izoenzymy są obecnie dobrze scharakteryzowane (**Srikhanta *et al*., 2013**). Asparaginaza typu I (EcAI, kodowana przez gen *ansA*) ulega ekspresji konstytutywnej, jest enzymem cytoplazmatycznym zaangażowanym w wykorzystywanie zgromadzonej w komórce asparaginy jako źródła azotu i charakteryzuje się niskim powinowactwem do L-Asn (mM K_M). Ekspresja asparaginazy typu II (EcAII, kodowana przez gen *ansB*) wymaga współzależnej aktywacji przez warunki beztlenowe (poprzez aktywator transkrypcji Fnr) i Crp (*ang. cyclic-AMP receptor protein*), a sam enzym występuje w periplazmie i ma wyższe powinowactwo względem L-Asn (μM K_M).

Struktura krystalograficzna EcAII została określona z rozdzielczością 2.3 Å (**Swain *et al*., 1993;** kod PDB: 3eca). Enzym ten jest homotetramerem, który należy traktować jako dimer dimerów ścisłych tworzonych przez łańcuchy białkowe A/B i C/D. Cztery podjednostki tetrameru są związane symetrią 222. Protomery (35.6 kDa) zbudowane są z dwóch domen αβ (większej N-terminalnej i mniejszej C-terminalnej) połączonych dwudziestodwu-aminokwasowym łącznikiem. Pomiędzy podjednostkami dimeru ścisłego na granicy styku domen C- i N-terminalnej należących do różnych protomerów znajdują się dwa miejsca aktywne, zatem w całej strukturze są cztery miejsca aktywne.

Cytoplazmatyczna asparaginaza z *E. coli* (EcAI) również składa się z czterech identycznych podjednostek tworzących dimer dimerów ścisłych. Jej protomery mają budowę podobną do protomerów EcAII (**Yun *et al*., 2007;** kody PDB: 2him, 2p2d, 2p2n), jednak organizacja tetrameru jest odmienna, kształtem przypominając obwarzanek. Enzym ten jest regulowany allosterycznie. W strukturze krystalicznej EcAI stwierdzono obecność czterech cząsteczek L-Asn w czterech równoważnych centrach allosterycznych i jednoczesną reorganizację czwartorzędowej struktury enzymu. Fakt ten sugeruje, że L-Asn jest nie tylko substratem EcAI, ale także efektorem

allosterycznym aktywującym enzym.

Asparaginazy typu bakteryjnego są dość zróżnicowaną grupą enzymów. W obrębie obu klas (I i II) mieszczą się zarówno enzymy pochodzenia eukariotycznego jak i prokariotycznego (**Borek i Jaskólski, 2001**), choć w jednym gatunku nie zawsze występują oba izoenzymy. Badania na drożdżach (*Saccharomyces cerevisiae*) wykazały obecność dwóch odrębnych asparaginaz (**Dunlop *et al.*, 1978**), jednak przeprowadzona analiza filogenetyczna sugeruje że wewnątrz- i zewnątrz-komórkowe enzymy u prokariotów i eukariotów wyewoluowały niezależnie (**Bonthon i Jaskólski, 1997**). Co więcej, asparaginazy należące do danej klasy mogą znacznie różnić się pod względem strukturalnym, jak np. EcAI i cytoplazmatyczna asparaginaza typu I z *Pyrococcus horikoshii* (**Yao *et al.*, 2005**). Z drugiej strony, na podstawie podobieństwa sekwencyjnego, do grupy asparaginaz typu bakteryjnego można również zaliczyć inne enzymy, jak np. ludzką lizofosfolipazę, której N-terminalna domena jest homologiczna do EcAI (**Karamitros i Konrad, 2014**). Należy również nadmienić, że oprócz aktwności asparaginazowej enzymy te mogą posiadać aktywność glutaminazową i zróżnicowany stopień powinowactwa do obu substratów (**Michalska i Jaskólski, 2006**).

Asparaginazy typu bakteryjnego są do dziś intensywnie badane (**Verma *et al.*, 2014; Karamitros i Konrad, 2014; Sudhir *et al.*, 2014; Pourhossein i Korbekandi, 2014; Tong *et al.*, 2014; Schalk *et al.*, 2014; Zuo *et al.*, 2015)**. W bazie PDB (*ang. Protein Data Bank*) znajduje się kilkadziesiąt struktur tych enzymów, w tym takie przykłady jak asparaginaza z *Erwinia chrysanthemi* określona z rozdzielczością 1Å (**Lubkowski *et al.*, 2003**). Olbrzymie zainteresowanie asparaginazami typu bakteryjnego, a szczególnie klasą II, jest spowodowane ich przeciwnowotworowymi właściwościami, które zostały odkryte w latach 60-tych. Obecnie istnieją trzy preparaty stosowane w ostrej białaczce limfoblastycznej (*ang. acute lymphoblastic leukemia,* ALL). Dwa z nich oparte są na enzymie EcAII (formy: natywna i PEG-asparaginaza), trzeci na enzymie z *Erwinia chrysanthemi* (**Andrade *et al.*, 2014**). Uważa się, że leki oparte na asparaginazach zatrzymują rozwój zmutowanych komórek białaczkowych, ponieważ podane pacjentom usuwają wolną L-asparaginę z krwioobiegu. Komórki białaczkowe, w przeciwieństwie do komórek zdrowych, które samodzielnie wytwarzają niezbędną do rozwoju L-asparaginę, uzależnione są od dostaw tego aminokwasu, więc po zastosowaniu leku "głodują". Niewykluczone jednak, że mechanizm działania leków jest bardziej skomplikowany (**Andrade *et al.*, 2014; Fine *et al.*, 2005; Iiboshi *et al.*, 1999; Suto *et al.*, 2010**). Poza potwierdzoną skutecznością przeciwnowotworową,

stosowane leki wywołują wiele efektów ubocznych, które są powodowane m.in. przez aktywność glutaminazową stosowanych enzymów oraz ich bakteryjne pochodzenie wywołujące reakcję immunologiczną. W związku z tym nadal doskonali się formę stosowanych leków i sposoby ich użycia (**Verma *et al.*, 2014; Fine *et al.*, 2005; Tong *et al.*, 2014**) oraz poszukuje się nowych asparaginaz mogących stanowić alternatywę dla obecnie używanych leków (**Sudhir *et. al.* 2014; Pourhossein i Korbekandi, 2014; Schalk *et al.*, 2014)**.

3.3.2. Asparaginazy typu roślinnego

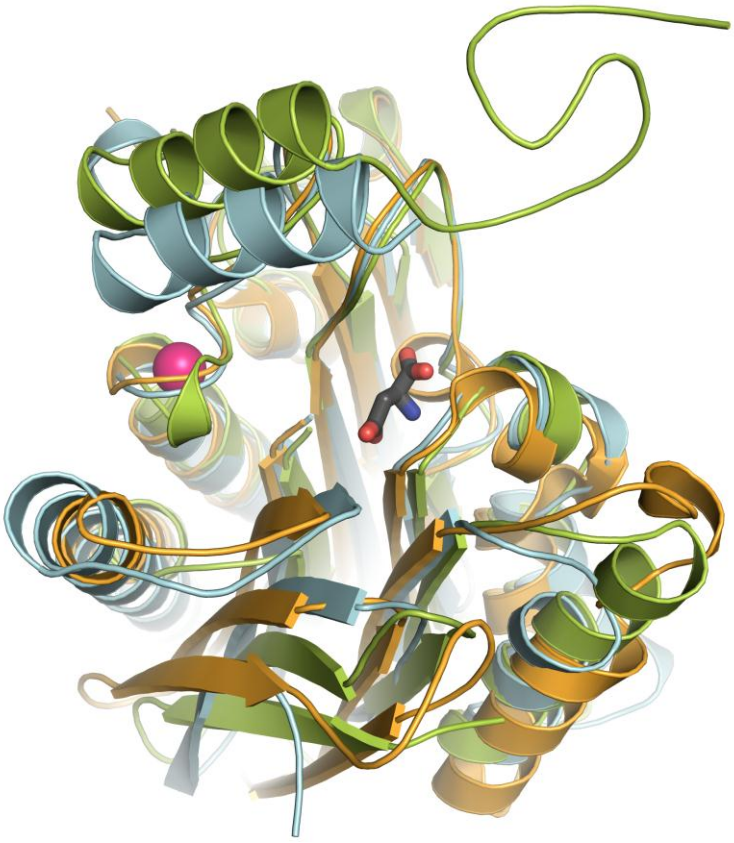
Druga grupa enzymów posiadających aktywność asparaginazową to asparaginazy typu roślinnego, nazywane również asparaginazami typu III. W grupie tej wyróżniamy enzymy potasowo-zależne i potasowo-niezależne. Asparaginazy typu roślinnego, ze względu na cechy strukturalne, zaliczane są równocześnie do rodziny Ntn-hydrolaz.

Zdolność hydrolizy L-Ans przez asparaginazy typu roślinnego jest znacznie niższa w porównaniu do asparaginaz typu bakteryjnego klasy II (wartości K_M mieszczą się w zakresie 1-30 mM, w zależności od organizmu źródłowego, zależności/niezależności od jonów K^+ i obecności jonów w testach kinetycznych). Z tego względu, jak dotąd, enzymy te nie zostały wykorzystane jako leki przeciwbiałaczkowe. Z drugiej strony, w ostatnich latach dostrzega się potencjał ludzkiego białka hASNase3 (jednego z przedstawicieli asparaginaz typu roślinnego), które jako lek nie powodowałoby reakcji immunologicznej.

3.3.2.1. Rodzina Ntn-hydrolaz

U wszystkich enzymów należących do rodziny Ntn-hydrolaz (*ang. N-terminal nucleophile*) rolę nukleofila pełni N-terminalna reszta aminokwasowa. Wszystkie Ntn-hydrolazy podlegają procesowi dojrzewania, który polega na uwolnieniu reszty nukleofila (treoniny, seryny lub cysteiny) w wyniku autoproteolitycznego przecięcia niefunkcjonalnego łańcucha polipeptydowego (**Branningan *et al.*, 1995**). Do rodziny tej należy wiele zróżnicowanych funkcjonalnie enzymów, dlatego jako przykłady podam (i scharakteryzuję) tylko te istotne dla wprowadzenia do niniejszej rozprawy: taspazę (Tas1) (**Khan *et al.*, 2005**), aspartyloglukozaminidazy (AGA) (**Oinonen *et al.*, 1995**) i asparaginazy typu roślinnego (**Michalska *et al.*, 2005, 2006, 2008-A; 2008-B**).

We wszystkich wymienionych białkach katalitycznym nukleofilem jest treonina, wszystkie są amidohydrolazami, ich substratami są pochodne L-asparaginianu. Asparaginazy typu roślinnego hydrolizują wiązanie amidowe łańcucha bocznego L-asparaginy oraz wiązanie β-peptydowe w peptydach izoaspartylowych (**Borek *et al.*, 2004**), AGA rozpoznaje L-asparaginę podstawioną w łańcuchu bocznym pierścieniem cukrowym, a taspaza jest proteazą katalizującą rozcinanie/potranslacyjne dojrzewanie białek jądrowych MLL (*ang. mixed-lineage leukaemia*).



Rysunek 3.3. Porównanie struktur trzech Ntn-hydrolaz. Nałożyłam atomy C_α pojedynczych αβ-heterodimerów modeli: asparaginazy z *Lupinus luteus* (kod PDB: 2gez - niebieskie łańcuchy białkowe z różową kulą reprezentującą jon sodu w pętli sodowej), ludzkiej AGA (kod PDB: 1apz - zielone łańcuchy z produktem reakcji - L-Asp w miejscu aktywnym przedstawionym na czarno z atomami O-czerwonymi, N-niebieskimi) i taspazy (kod PDB: 2a8j - pomarańczowe łańcuchy białkowe).

Znaczne podobieństwo sekwencyjne tych enzymów (ok. 60%) idzie w parze z podobieństwem strukturalnym (r.m.s.d. pozycji C_α ok. 1.9 Å) (Rysunek 3.3). W dojrzałej formie są heterotetramerami, lub raczej dimerami αβ-heterodimerów powstających podczas procesu autokatalitycznego rozcięcia prekursorowego łańcucha na podjednostki α i β. Heterodimery αβ (taspazy, aspartyloglukozaminidaz i asparaginaz typu roślinnego) posiadają wspólny typ pofałdowania, a dokładniej zbudowane są z dwóch arkuszy β oflankowanych z obu stron helisami α, tworząc tzw. "kanapkę αββα". Cechą szczególną asparaginaz typu roślinnego jest obecność pętli sodowej (koordynującej jon sodu), która za pomocą wiązania wodorowego oddziałuje z katalityczną treoniną. Pętla ta występuje również w taspazie, jednak zawiera ona cząsteczkę wody zamiast jonu metalu, prawdopodobnie błędnie wymodelowaną

(**Michalska i Jaskólski, 2006**). W AGA analogiczna pętla nie koordynuje jonów metalu lecz jest stabilizowana przez mostek dwusiarczkowy (**Oinonen *et al.*, 1995**).

3.3.2.2. Enzymy potasowo-zależne i potasowo-niezależne

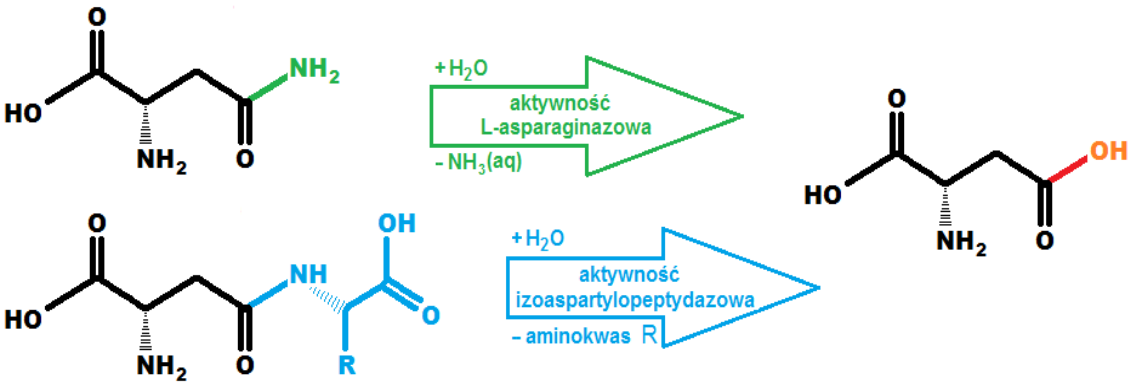
W grupie asparaginaz typu roślinnego wyróżnia się dwie podgrupy: enzymy potasowo-niezależne i potasowo-zależne (**Sodek *et al.*, 1980**). Znajdujące się w środowisku kationy metali jednowartościowych nie mają wpływu na aktywność L-asparaginazową enzymów K^+-niezależnych, natomiast regulują tą aktywność u enzymów K^+-zależnych. Badania biochemiczne enzymów z *Arabidopsis thaliana* (**Bruneau *et al.*, 2006**) oraz *Lotus japonicus* (**Credali *et al.*, 2011**) wykazały, że aktywność L-asparaginazowa potasowo-zależnych białek jest największa w obecności jonu K^+ i zmniejsza się wraz z zastosowaniem innych jonów według uogólnionego schematu: $K^+ > Na^+ \approx Rb^+ > Li^+ \approx Cs^+$. Do czasu podjętych przeze mnie badań molekularne podstawy tego zjawiska nie były znane oraz nie określono żadnej struktury krystalograficznej enzymu potasowo-zależnego. Dla K^+-zależnej asparaginazy z *Lotus japonicus* zaproponowano jedynie model struktury przestrzennej uzyskany przy pomocy modelowania molekularnego i na jego podstawie wytypowano trzy reszty (Glu248, Asp285 i Glu286), które miałyby koordynować jon potasu i odpowiadać za aktywność enzymu (**Credali *et al.*, 2011**).

Asparaginazy potasowo-niezależne to podgrupa bardzo dobrze poznana i reprezentowana w bazie PDB przez kilkanaście struktur rentgenowskich. Zalicza się do niej nie tylko enzymy pochodzenia roślinnego (klasyfikowane na podstawie badań biochemicznych i podobieństwa sekwencyjnego), ale także (na podstawie podobieństwa strukturalnego) homologi tych białek z innych organizmów. Określono dotąd struktury następujących enzymów: LlA - roślinnej asparaginazy z *Lupinus luteus* (kod PDB: 2gez, **Michalska *et al.*, 2006**), bakteryjnego EcAIII z *Escherichia coli* (kody PDB: 1t3m, 2zal, 1jn9, 1k2x, 2zak, 3c17; **Prahl *et al.*, 2004; Michalska et al., 2005, 2008-A; 2008-B)**, ludzkiego hASNase3 (kody PDB: 4gdt, 4gdu, 4gdv, 4gdw, 4hlo, 4hlp, 4o0c, 4o0d, 4o0e, 4o0f, 4o0g, 4o0h; **Nomme *et al.*, 2012, 2014; Su *et al.*, 2013**) i pochodzącą ze świnki morskiej gpASNase3 (kody PDB: 4o47, 4o48; **Schalk i Lavie, 2014**).

3.3.2.3. Aktywność L-asparaginazowa i izoaspartylopeptydazowa

Asparaginazy typu roślinnego katalizują dwa rodzaje reakcji (Rysunek 3.4). Aktywność L-asparaginazowa umożliwia hydrolizę L-asparaginy z uwolnieniem amoniaku, który wykorzystywany jest jako źródło azotu w trakcie biosyntezy białek. Aktywność ta jest szczególnie ważna w szybko rozwijających się tkankach i podczas syntezy białek zapasowych w dojrzewających nasionach. Aktywność izoaspartylopeptydazowa pozwala eliminować nieprawidłowe i szkodliwe peptydy izoaspartylowe powstające spontanicznie i akumulujące się w białkach zapasowych nasion.

Rysunek 3.4. Aktywności enzymatyczne asparaginaz typu roślinnego.



Różne asparaginazy typu roślinnego mają zróżnicowane preferencje substratowe, które nie mają związku z podziałem tej grupy na enzymy potasowo-niezależne i potasowo-zależne. Badania biochemiczne *in vitro* wykazały dominującą aktywność wobec β-peptydów i tylko drugorzędną aktywność wobec L-Asn u K^+-niezależnych asparaginaz z *L. luteus* (LlA), *E. coli* (EcAIII) i *L. japonicus* (LjNSE2) (**Borek *et al.*, 2004; Credali *et al.*, 2011**). Z kolei K^+-zależny enzym z *L. japonicus* (LjNSE1) przejawia modulowaną aktywność L-asparaginazową i brak aktywności izoaspartylopeptydazowej (**Credali *et al.*, 2011**). W przypadku *A. thaliana* dostępne dane literaturowe są niespójne. Wcześniejsze badania, świadczą o niewielkiej przewadze aktywności L-asparaginazowej enzymu potasowo-niezależnego i absolutnej specyficzności wobec L-asparaginy enzymu potasowo-zależnego, który β-peptydów nie hydrolizuje w ogóle (**Bruneau *et al.*, 2006**). Późniejsze wyniki sugerowały z kolei, że oba enzymy (K^+-niezależny i K^+-zależny) hydrolizują oba rodzaje substratów, ale K^+-zależne białko zdecydowanie preferuje L-Asn (**Gabriel *et al.*, 2012**). W przypadku

ssaczych homologów asparaginazy typu roślinnego (hASNase3 i gpASNase3) badania dowodzą, że enzymy te katalizują obie reakcje, przy czym hASNase3 spośród wszystkich testowanych substratów wykazuje najwyższą aktywność wobec cząsteczki estru β-metylowego L-Asp (**Cantor *et al.*, 2009**).

3.3.3. Asparaginazy typu *Rhizobium etli*

Do grupy asparaginaz typu *Rhizobium etli* początkowo zaklasyfikowano zaledwie osiem enzymów (**Borek i Jaskólski, 2001**), jednak dzisiaj do przedstawicieli tej grupy zalicza się ok. 350 białek, w większości pochodzących z prokariontów (konkretnie z mikroorganizmów glebowych i morskich) z zaledwie kilkoma przykładami z eukariontów (grzybów i mchów). Analiza sekwencji enzymów tej grupy wykazała obecność czterech wysoce zachowawczych motywów: RSx(2)KPxQA, ALxCASH, NCSGKHxGxL, AMx(3)Px(2)VAGxGRx(2)TxLM, oraz jednoczesną znaczną odmienność sekwencyjną tej grupy w odniesieniu do asparaginaz typu bakteryjnego i roślinnego (**Moreno-Enriquez *et al.*, 2012**).

Rhizobium etli, od którego grupa asparaginaz otrzymała swą nazwę, to gatunek bakterii brodawkowych wchodzących w symbiozę z fasolą zwyczajną (*Phaseolus vulgaris*) wyróżniony w 1993 roku (**Segovia *et al.*, 1993**). W *R. etli* zidentyfikowane zostały dwie asparaginazy, których obecność potwierdzono w ekstraktach komórkowych szczepu dzikiego, nie określając precyzyjnie poziomu aktywności tych enzymów (**Huerta-Zepeda *et al.*, 1997**). Asparaginaza I jest enzymem termostabilnym o ekspresji konstytutywnej. Asparaginaza II to enzym termolabilny, którego ekspresja jest indukowana przez L-asparaginę. Sekwencja asparaginazy I nie jest znana, natomiast o genie asparaginazy II (*ansA*) wiadomo, że znajduje się w operonie ansRPAB (**Ortuno-Olea i Duran-Vargas, 2000**), który zlokalizowano na niezwykle stabilnym plazmidzie p42e, uznanym za chromosom poboczny (*ang. secondary chromosome*) (**Landeta *et al.*, 2011**). Dzięki aktywacji dwóch enzymów z tego operonu: asparaginazy II i aspartazy, *R. etli* może wykorzystywać egzogenną L-asparaginę jako źródło węgla i azotu. Co więcej wysoki poziom obu enzymów w bakteroidach *R. etli* (a także *Rhizobium lupini*) może świadczyć o ich znaczeniu dla symbiozy (**Kretovich *et al.*, 1981; Huerta-Zepeda *et al.*, 1996**).

Asparaginaza II została niedawno scharakteryzowana pod względem biochemicznym (**Moreno-Enriquez *et al.*, 2012**). Enzym, otrzymany i oczyszczony

metodami biotechnologicznymi, zawierający N-terminalny peptyd fuzyjny, jest homodimerem o masie 102 kDa. Asparaginaza II z *R. etli* jest specyficzna wobec L-asparaginy, a poziom efektywności hydrolizy tego substratu jest zbliżony do poziomu wykazanego dla enzymu z *Erwinia chrysantemi* (**Kotzia i Labrou, 2007**). Aktywność glutaminazowa nie została wykryta. Optymalne warunki działania enzymu wynoszą pH 9.0 i 37ºC. W środowisku zasadowym zmiany temperatury nie mają znaczącego wpływu na aktywność enzymu. W środowisku obojętnym (pH 7.5) w 37ºC aktywność enzymu jest trzy razy mniejsza i rośnie wraz ze wzrostem temperatury w zakresie 25 – 50ºC, osiągając maksymalną wartość równą połowie poziomu zarejestrowanego w warunkach zasadowych.

Jak dotąd w bazie PDB nie ma żadnego przedstawiciela aspaginaz typu *R. etli*. Stworzony *in silico* model asparaginazy II z *Rhizobium etli* sugeruje pofałdowanie zupełnie inne niż to znane z asparaginaz typu bakteryjnego (**Huerta-Saquero *et al.*, 2013**). Co więcej, odmienność sekwencyjna i strukturalna wskazuje na istnienie odrębnego profilu immunogenetycznego enzymów tego typu. Podsumowując, dostępne informacje na temat asparaginazy II z *Rhizobium etli* jednoznacznie wskazują, że enzym ten mógłby stanowić podstawę nowoczesnego leku na białaczkę (**Huerta-Saquero *et al.*, 2013**).

4. Materiały i metody

4.1. Systemy ekspresyjne

W procesie określania struktury przestrzennej białek metodą rentgenografii, niezbędne jest uzyskanie kryształu białka, co z kolei wymaga znacznych ilości białka o wysokiej czystości. Stąd w biokrystalografii wysoce użytecznym narzędziem są systemy ekspresyjne, które pozwalają na stosunkowo szybką, tanią i wydajną produkcję białek rekombinowanych. Wykorzystywane w niniejszej pracy systemy ekspresyjne składają się z plazmidowych wektorów ekspresyjnych zawierających informację genetyczną dotyczącą badanego białka (tj. sekwencję kodującą w postaci tzw. insertu DNA) i bakteryjnego gospodarza (*Escherichia coli*), w którym następuje ekspresja białka. Przygotowanie systemu ekspresyjnego dla danego białka wymaga zastosowania różnorodnych metod i materiałów.

4.1.1. Amplifikacja i analiza DNA

Dowolny fragment DNA można powielić *in vitro* przeprowadzając reakcję PCR (*ang. Polymerase Chain Reaction*) (**Mullis *et al.,* 1986**). Standardowa reakcja PCR zawiera cyklicznie powtarzane trzy etapy: (I) rozplecenie matrycowego DNA w temperaturze 95ºC, (II) hybrydyzacja starterów do matrycy w temperaturze ich topnienia, (III) elongacja starterów zgodna z sekwencją matrycowej nici DNA przeprowadzana przez polimerazę DNA w temperaturze ok. 72ºC oraz (IV) etap końcowy: dodatkowej elongacji niedokończonych jednoniciowych fragmentów powielonych cząsteczek DNA.

W pracy doświadczalnej reakcję tą wykorzystywałam, aby wyłowić pożądany gen z genomowego DNA (tworząc insert łączony następnie z wektorem ekspresyjnym), zweryfikować poprawność konstrukcji tworzonych wektorów ekspresyjnych oraz do wytworzenia wektorów metodą PIPE. Używałam trzech różnych polimeraz syntetyzujących jednoniciowy DNA w kierunku od 5’ do 3’: polimerazy Taq (Fermentas) oraz polimeraz PWO (Roche) i Pfu (Thermo Scientific), które posiadają dodatkową aktywność 3’ → 5’ egzonukleazy. Produkty reakcji PCR analizowałam przeprowadzając elektroforezę w żelach agarozowych (**Sambrook i Russell, 2001**).

4.1.2. Plazmidowe wektory ekspresyjne

Plazmidy to koliste cząsteczki dwuniciowego DNA, naturalnie występujące w cytoplazmie komórek bakteryjnych, które dzięki obecności sekwencji *ori* wykazują zdolność do samodzielnej replikacji niezależnie od chromosomu bakteryjnego. Na podstawie plazmidów stworzone zostały komercyjne wektory wykorzystywane jako nośniki heterologicznej informacji genetycznej, która za ich pomocą wprowadzona zostaje do komórek bakteryjnych. Odpowiednia konstrukcja wektorów umożliwia transkrypcję i translację tej informacji przy wykorzystaniu aparatu enzymatycznego komórki gospodarza. Wszystkie wektory posiadają także markery selekcyjne, umożliwiając wzrost jedynie tym bakteriom, które posiadają dany wektor.

Obecnie istnieje wiele rodzajów wektorów, które oprócz transportowanego genu niosą cały wachlarz przeróżnych cech dodatkowych (**Sorensen i Mortensen, 2005**). W swojej pracy badawczej korzystałam z dwóch komercyjnych wektorów: pET151/D-topo (Invitrogen) oraz pMALc5x (NEB), na podstawie których stworzyłam trzy rodzaje plazmidowych wektorów ekspresyjnych, trzema różnymi sposobami klonowania. Wszystkie używane przeze mnie wektory posiadały gen kodujący β-laktamazę (*bla*) warunkujący oporność na antybiotyk ampicylinę, który dodawałam do pożywki hodowlanej.

4.1.2.1. Wektor pMALc5x i klasyczna metoda klonowania

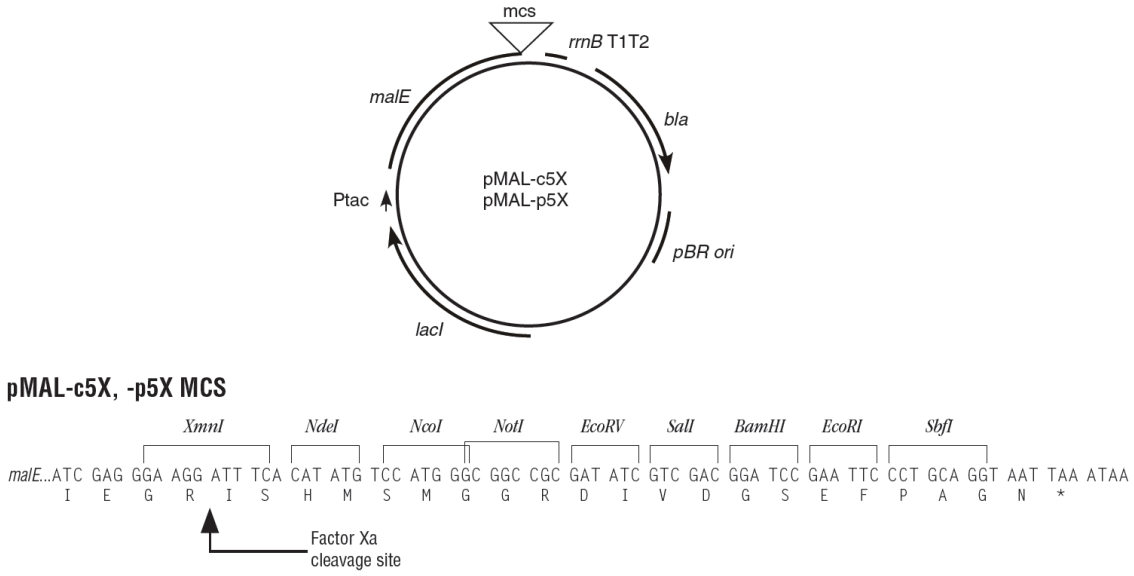
Do wektora pMALc5x (Rysunek 4.1) obcy gen wprowadzany jest w rejonie tzw. miejsca wielokrotnego klonowania (*ang. multiple cloning site,* mcs) najpowszechniejszą, klasyczną trzyetapową metodą. Pierwszy etap to reakcja PCR podczas której przygotowywany jest insert. Drugi etap obejmuje reakcje enzymatyczne polegające na cięciu (tymi samymi) endonukleazami restrykcyjnymi wektora oraz insertu, dzięki czemu powstałe w ten sposób fragmenty DNA mają komplementarne lepkie końce. Podczas prowadzenia doświadczeń korzystałam z dwóch enzymów restrykcyjnych: SbfI (Fermentas) i XmnI (Promega). Trzeci etap to reakcja łączenia tych fragmentów, podczas której pomiędzy zasadami azotowymi komplementarnych nukleotydów spontanicznie tworzą się wiązania wodorowe. Natomiast wiązania fosfodiestrowe, pomiędzy końcem 3’ hydroksylowym, a końcem 5’ fosforanowym powstają dzięki zastosowaniu ligazy. Podczas doświadczeń wykorzystywałam T4 DNA

ligazę (Fermentas), która katalizuje reakcję zależnie od adenozynotrójfosforanu (ATP).

Wprowadzona do pMALc5x w obszarze *mcs* heterologiczna sekwencja kodująca znajduje się pod kontrolą silnego promotora transkrypcji *tac*, hybrydy promotorów *trp* i *lac*UV5, który zawiera sekwencję operatora laktozowego. Wektor pMALc5x posiada gen kodujący białko represorowe operonu laktozowego, tj. represor Lac (*lacI*). Gdy represor Lac wiąże się do promotora *tac*, a konkretnie do operatora laktozowego, uniemożliwia pochodzącej z bakteryjnego gospodarza polimerazie RNA syntezę mRNA. Ekspresję obcego genu, a właściwie jego transkrypcję, indukuje się za pomocą IPTG (analog laktozy), który wiąże białko represorowe Lac i odblokowuje promotor *tac* umożliwiając polimerazie RNA syntezę mRNA.

Wektor pMALc5x posiada również gen kodujący białko wiążące maltozę (MBP, *ang. Maltose Binding Protein*) (*malE*). Gen ten, razem z wprowadzonym podczas klonowania obcym genem, znajduje się pod kontrolą promotora *tac* i wraz z nim ulega transkrypcji oraz translacji. Po ekspresji oba białka można rozdzielić dzięki istniejącej pomiędzy nimi sekwencji rozpoznawanej przez czynnik Xa.

Rysunek 4.1. Schemat przedstawiający cechy wektora pMALc5x. Rysunek zapożyczony z instrukcji: pMAL protein fusion and purification system (NEB).

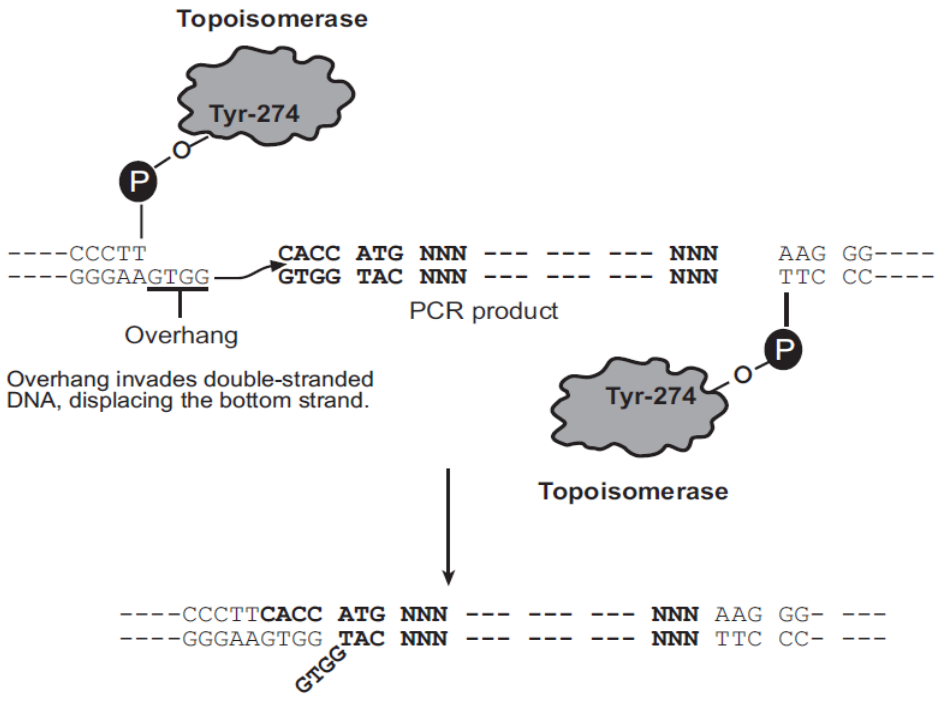


4.1.2.2. Wektor pET151/D-topo i system TOPO

Wprowadzenie pożądanej sekwencji kodującej białko do wektora pET151/D-topo odbywa się dwuetapowo. Pierwszym etapem jest przygotowanie insertu DNA, który powstaje w wyniku reakcji PCR, podczas której odpowiednio zaprojektowany starter

dołącza do 5'-końca genu dodatkową sekwencję CACC. Drugim etapem jest ligacja insertu i wektora (Rysunek 4.2). Komercyjny wektor pET151/D-topo dostępny jest jako nić DNA (a nie typowa kolista zamknięta cząsteczka), której jeden koniec jest tępy a drugi lepki, zawierający sekwencję GTGG, komplementarną do sekwencji CACC insertu. Dzięki takiej konstrukcji końców obu cząsteczek mogą one połączyć się w odpowiedniej orientacji spontanicznie dzięki parowaniu się komplementarnych nukleotydów. Natomiast wytworzenie wiązania fosfodiestrowego pomiędzy cząsteczkami deoksyryboz, następuje dzięki obecności enzymu, topoizomerazy, a dokładniej dzięki obecności wiązania pomiędzy topoizomerazą i wektorem (wiązanie kowalencyjne pomiędzy grupą 3'-fosforanową nici DNA i resztą tyrozyny enzymu). Wiązanie to zostaje zaatakowane przez grupę 5'hydroksylową nici insertu, a energia z przerwanego wiązania wektor - enzym (uwolnienie topoizomerazy) jest zachowana w postaci nowo utworzonego wiązania kowalencyjnego pomiędzy cząsteczkami DNA (złączenie inseru i wektora) (**Shuman, 1994**).

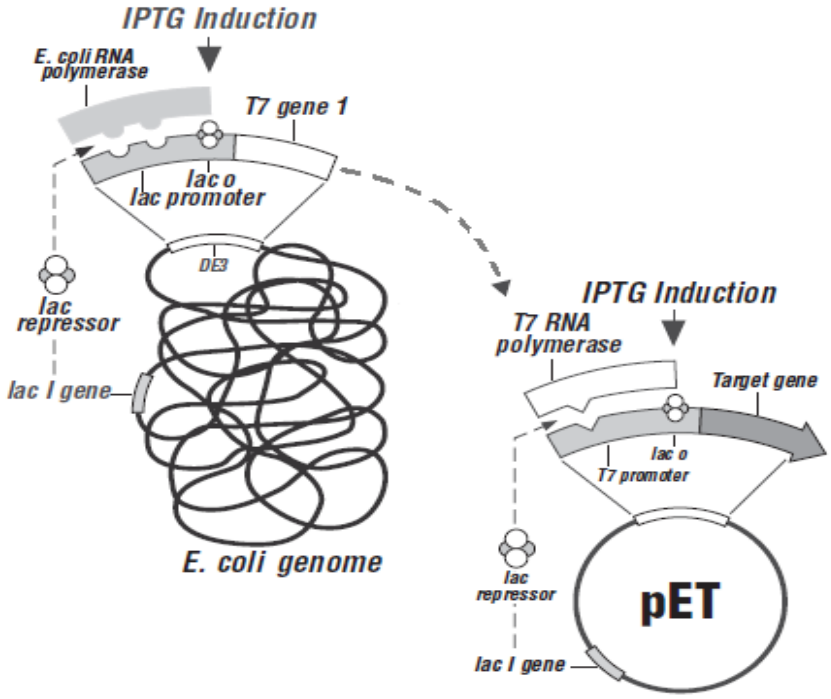
Rysunek 4.2. Ligacja insertu i wektora pET151/D-topo. Rysunek zapożyczony z instrukcji Champion pET Directional TOPO Expression Kits (Invitrogen).



Wprowadzony do pET151/D-topo heterologiczny gen znajduje się pod kontrolą promotora T7, zawierającego sekwencję operatora laktozowego (*lac O*) właściwego dla polimerazy RNA z faga T7 (Rysunek 4.3). Wektor pET151/D-topo posiada również gen kodujący białko represorowe operonu laktozowego (*lacI*). Represor Lac w wektorach

typu pET zapewnia podwójną kontrolę ekspresji. Po pierwsze blokując transkrypcję wklonowanego obcego genu z promotora T7. Po drugie zapobiegając ekspresji samej polimerazy T7, której gen znajduje się pod kontrolą promotora *lac*UV5 w genomie bakteryjnym komórek ekspresyjnych posiadających tzw. fragment λDE3 pochodzący z genomu faga λ. Ekspresja polimerazy T7 jak i genu docelowego następuje w obecności induktora IPTG.

Rysunek 4.3. Układ ekspresyjny wektorów typu pET wykorzystujących T7 polimerazę RNA. Schemat zapożyczony z instrukcji: pET system manual (Novagen).



Wektor pET151/D-topo wprowadza na N-końcu białka docelowego fragment łańcucha polipeptydowego (4 kDa), który zawiera znacznik histydynowy (sekwencję sześciu reszt histydyn) oraz miejsce rozpoznawane przez proteazę TEV.

4.1.3. Metoda PIPE

Metoda PIPE (*ang. Polymerase Incomplete Primer Extension*) jest sposobem manipulacji fragmentami DNA przy użyciu reakcji PCR. Może ona służyć zarówno do klonowania jak i mutagenezy **(Klock i Lesley, 2009)**. W swojej pracy badawczej wykorzystałam ją do modyfikacji wektorów ekspresyjnych, polegającej na usunięciu fragmentu sekwencji nukleotydowej kodującej jeden z elementów tandemowej konstrukcji eksprymowanego białka (Rozdział 5.2.2.3).

Metoda opiera się na kilku prostych założeniach. (I) Podczas reakcji PCR powstaje mieszanina fragmentów DNA posiadających 5' jednoniciowe końce różnej długości, szczególnie gdy pominie się ostatni etap reakcji, tj. dodatkową, wypełniającą te końce elongację (Rozdział 4.1.1). (II) Odpowiednio skonstruowane startery do PCR spowodują, że te końcowe jednoniciowe fragmenty DNA będą miały precyzyjnie określoną sekwencję i będą spontanicznie łączyć się z komplementarnymi jednoniciowymi fragmentami DNA w wyniku parowania się nukleotydów. (III) Tak połączone w kolistą cząsteczkę DNA produkty PCR (bez wytwarzania wiązania fosfodiestrowego) można potraktować jak wektor i przeprowadzić nim transformację komórek, w których, podczas replikacji, enzymy bakteryjnego gospodarza dokończą dzieła klonowania. Sam proces namnażania świeżo przygotowanych wektorów w nieekspresyjnych komórkach (jak również następująca po tym weryfikacja i izolacja wektorów) jest standardowym etapem tworzenia wektorów ekspresyjnych posiadających poprawną konstrukcję i sekwencję (Rozdział 5.2.2.4)

4.1.4. Ekspresja białek

Aby umożliwić ekspresję rekombinowanych białek z przygotowanych wektorów ekspresyjnych, wykonuje się transformację chemiczną kompetentnych komórek ekspresyjnych, poddając je np. szokowi termicznemu (*ang. heat shock*). Następnie komórki te hoduje się na pożywkach z antybiotykiem, aż do osiągnięcia logarytmicznej fazy wzrostu ($OD_{600} = 0.8$) i indukuje się ekspresję dodając IPTG.

Do ekspresji używałam pięciu różnych rodzajów komórek ekspresyjnych (Tabela 4.1) oraz trzech różnych pożywek. Pożywki LB i TB przygotowywałam według standardowych przepisów (**Sambrook i Russell, 2001**). Pożywka minimalna składała się z syntetycznej pożywki M9 (**Sambrook i Russell, 2001**) uzupełnionej glukozą, witaminami i aminokwasami. Pożywkę tą przygotowywałam w oparciu o komercyjny zestaw SelenoMethionine Medium Base & Nutrient Mix (MDL), który zasadniczo stworzony został do ekspresji białek znakowanych selenomietioniną (Se-Met). Produkując białka znakowane do pożywki dodawałam roztwór Se-Met, natomiast gdy prowadziłam ekspresję białek nieznakowanych zwyczajnie używałam pożywki minimalnej bez tego aminokwasu.

Ponadto w trakcie ekspresji stosowałam różne temperatury i różne stężenia IPTG. Ekspresję białek rekombinowanych w fuzji z MBP prowadziłam zawsze w pożywce z

dodatkiem glukozy. Dodatek ten blokował ekspresję amylazy (gen obecny na chromosomie *E. coli*), która podczas oczyszczania mogłaby rozłożyć amylozowe złoże stosowane w chromatografii powinowactwa.

Tabela 4.1. Wybrane cechy komórek stosowanych podczas ekspresji białek rekombinowanych.

nazwa komórek (producent)	najważniejsze wybrane cechy komórek ekspresyjnych
BL21 star (Invitrogen)	- obecny fragment λDE3 umożliwiający ekspresję T7 polimerazy RNA, znajdującej się pod kontrolą *lac*UV5, za pomocą IPTG - nieaktywne proteazy (*lon* i OmpT), co obniża poziom degradacji eksprymowanych heterologicznych białek - nieaktywny gen RNazy E (*rne131*), co powoduje wzrost stabilności mRNA
NEB express (NEB)	- nieaktywne proteazy (*lon* i OmpT)
OverExpress C41 i C43 (Lucigen)	- obecny fragment λDE3 umożliwiający ekspresję T7 polimerazy RNA - nieaktywne proteazy (*lon* i OmpT) - szczepy te posiadają nieokreślone mutacje, które powodują efektywną ekspresję białek toksycznych (**Miroux i Walker, 1996**); prawdopodobnie zawierają wewnętrzne błony dzielące wnętrze komórki na dodatkowe przedziały
BL21 magic (MCSG)	- obecny fragment λDE3 umożliwiający ekspresję T7 polimerazy RNA - plazmid „magic" kodujący rzadko wykorzystywane przez bakterie tRNA argininowe (rozpoznające kodony: AGG i AGA) oraz izoleucynowe (ATA), zawierający również gen oporności na kanamycynę

4.2. Oczyszczanie białek

Wyprodukowane w *E. coli* rekombinowane białka uwalnia się (liza komórkowa), a następnie poddaje procesowi oczyszczania, który pozwala pozbyć się wszelkich niepożądanych elementów komórkowych (jak błony komórkowe, kwasy nukleinowe, węglowodany, naturalne białka *E. coli*, itp.) i uzyskać homogenny, nadający się do krystalizacji preparat docelowego białka. Najczęściej oczyszcza się białka rozpuszczalne, jednak czasem korzysta się również z niepoprawnie pofałdowanych agregatów, pochodzących z ciałek inkluzyjnych, doprowadzając do ich rozplecenia (za pomocą czynników chaotropowych), a następnie renaturacji (Rozdział 5.2.4.1).

Istnieje wiele metod stosowanych podczas oczyszczania, które opierają się na przeróżnych właściwościach fizykochemicznych białek, stąd wymieniam tylko te, które stosowałam w trakcie swoich prac badawczych. Chromatografia jonowymienna wykorzystuje różnice w ładunkach cząsteczek białek, a filtracja żelowa różnice w promieniach hydrodynamicznych separowanych białek. Chromatografia powinowactwa opiera się na specyficznym oddziaływaniu pomiędzy złożem a rekombinowanym białkiem, które jest efektem modyfikacji wprowadzonej do białka metodami inżynierii

genetycznej na etapie klonowania. Wykonywałam np. chromatografię na złożu amylozowym, które wiąże białko MBP (białko wiążące maltozę, *ang. Maltose Binding Protein)* oraz chromatografię na złożu niklowym (chelatującym jony Ni^{2+}), które wiąże znacznik histydynowy. Krótkie peptydowe znaczniki oraz białka fuzyjne usuwane są za pomocą proteaz (np. TEV lub Xa). Rozdziały chromatograficzne mogą być wykonywane tradycyjną metodą grawitacyjną (niekiedy przyspieszaną przy użyciu podciśnienia) lub w systemie FPLC (wysokosprawnej chromatografii cieczowej, *ang. Fast Performance Liquid Chromatography*) z zastosowaniem nadciścienia.

Podczas oczyszczania białek oprócz metod chromatograficznych stosowałam także techniki wykorzystujące zjawisko wysalania białek (wytrącania ich z roztworu) oraz osmozy (dializa). Używane przeze mnie roztwory poza składnikami podstawowymi (bufor i sól) często zawierały składniki dodatkowe (czynniki redukcyjne, detergenty, itp.), za pomocą których próbowałam wpłynąć na właściwości fizyko-chemiczne oczyszczanych obiektów **(Bondos i Bicknell, 2003; Prive, 2007).**

4.3. Weryfikacja jakości preparatu białka

Podstawową stosowaną przeze mnie metodą weryfikacji jakości preparatów białka była elektroforeza w żelu poliakrylamidowym w warunkach denaturujących (SDS-PAGE). W metodzie tej, SDS rozwija łańcuch polipeptydowy i nadaje mu jednolity ładunek ujemny proporcjonalny do wielkości danej cząsteczki. Rozdział zdenaturowanych naładowanych cząsteczek odbywa się w polu elektrycznym, a szybkość migracji przez pory żelu uzależniona zostaje od ich masy. Dzięki użyciu odpowiedniego markera mas rozdział białek względem wielkości, a dokładniej rozdział podjednostek białkowych (składających się z pojedynczych łańcuchów), umożliwia oszacowanie wielkości białek/podjednostek znajdujących się w badanej próbce i dokonanie na tej podstawie ich wstępnej identyfikacji, z jednoczesnym określeniem czystości preparatu.

Do oszacowania stężenia białka w uzyskanych próbkach wykorzystywałam metodę spektrofotometryczną i spektrofotometr Nano Drop (Thermo Scientific). Metoda ta polega na pomiarze absorpcji światła o długości fali $\lambda = 280$ nm przez reszty aromatyczne tyrozyny i tryptofanu dla białka o znanej sekwencji. Do obliczenia stężenia białka stosuje się prawo Lamberta-Beera.

Kolejną metodą, z której wyników korzystałam jest spektrometria mas, przy czym

wykonanie doświadczeń zleciłam Pracowni Spektrometrii Mas IChB PAN. Metoda ta pozwala z dużą dokładnością określić masę cząsteczkową białka zawartego w preparacie. Umożliwia to między innymi: (I) potwierdzenie tożsamości białka, (II) wykrycie uszkodzeń cząsteczki spowodowanych przez proteazy oraz (III) weryfikację obecności atomów ciężkich w białku (np. selenu). Ponadto spektrometria mas umożliwia identyfikację białek na podstawie wzoru peptydów, uzyskanych podczas kontrolowanej proteolizy.

Białka poddawałam również testom na analizatorze Zetasizer Nano (Malvern Instruments Ltd.) za pomocą którego określałam wartości promienia hydrodynamicznego cząstek zawartych w roztworze oraz stopień ich dyspersji. Instrument ten wykorzystuje zjawisko dynamicznego rozpraszania światła (*ang. Dynamic Light Scattering*, DLS) do śledzenia ruchów Browna cząstek w roztworze. W metodzie tej wielkość cząstek jest obliczana z równania Stokes'a- Einstein'a, które określa zależność między wielkością cząstki i jej prędkością na skutek ruchów Browna. Jednak (według mnie), podawane przez analizator promienie cząstek są tylko wartościami przybliżonymi, gdyż założenie o idealnej sferyczności badanych białek jest rzadko spełnione. Ponadto pomiary analizatora DLS obarczone są sporym błędem ze względu na wpływ lepkości roztworów na ruchy znajdujących się w nim cząstek (nie wykonuje się kalibracji instrumentu, a do obliczeń podaje się chemiczny skład roztworów nie uwzględniający stężenia białka). Instrument prowadzi obliczenia dwoma różnymi algorytmami (**Malvern Instruments Ltd., 2013)** podając dwa często znacznie różniące się wyniki. Pierwszy z nich, podający uśrednioną wartość promienia hydrodynamicznego jest wiarygodny tylko i wyłącznie wtedy gdy w roztworze obecne są cząstki jednego rodzaju (Hydrodynamic Radius). Drugi (Distribution Results) podje ten sam parametr dla poszczególnych grup obiektów o różnej wielkości.

4.4. Krystalizacja białek

Do krystalizacji białek zastosowałam metodę dyfuzji par, w której siłą napędową krystalizacji jest dążenie zamkniętego układu do osiągnięcia stanu równowagi. Taki układ krystalizacyjny tworzą rezerwuar oraz kropla będąca mieszaniną preparatu białka i roztworu rezerwuarowego. Dyfuzja składników lotnych pomiędzy kroplą a rezerwuarem, prowadząca do wyrównania stężeń czynnika strącającego w obu roztworach, powoduje zagęszczenie białka w kropli i wzrost kryształów.

Tworzenie się kryształów danego białka uwarunkowane jest szeregiem indywidualnych czynników stanowiących warunki krystalizacji (**Bergfors, 1999**). Aby takie warunki wyznaczyć prowadziłam krystalizacje przesiewowe, głównie przy użyciu robota krystalizacyjnego Gryphon (Art Robbins Instruments) i komercyjnych zestawów do krystalizacji (Tabela 4.2). Użycie robota pozwalało przetestować jednocześnie, na jednej płytce krystalizacyjnej, 96 roztworów/warunków krystalizacyjnych, przy minimalnym zużyciu preparatu białka, ponieważ nastawiane krople siedzące miały całkowite objętości w zakresie 0.2 - 0.6 μL.

Tabela 4.2. Zestawienie używanych komercyjnych zestawów roztworów do krystalizacji.

producent	numer katalogowy	nazwa	liczba roztworów
MDL	MD1-01	Structure Screen 1	50
MDL	MD1-02	Structure Screen 2	50
MDL	MD1-14	Clear Strategy Screen I	24
MDL	MD1-15	Clear Strategy Screen II	24
MDL	MD1-29	PACT premier Screen	96
MDL	MD1-37	JCSG plus Screen	96
MDL	MD1-46	Morpheus Screen	96
MDL	MD1-50	PGA Screen	96
MDL	MD1-59	MIDAS Screen	96
Hampton Research	HR2-110	Crystal Screen	50
Hampton Research	HR2-112	Crystal Screen 2	48
Hampton Research	HR2-114	MembFac Screen	48
Hampton Research	HR2-126	PEG/Ion Screen	48
Hampton Research	HR2-098	PEG/Ion 2 Screen	48
Hampton Research	HR2-211	Grid Screen Ammonium Sulfate	24
Hampton Research	HR2-213	Grid Screen PEG 6000	24
Hampton Research	HR2-247	Grid Screen Sodium Malonate	24
Hampton Research	HR2-215	Grid Screen MPD	24
Hampton Research	HR2-138	Additive Screen HT	96
Hampton Research	HR2-134	Index HT Screen	96
Hampton Research	HR2-137	MembFac HT	96
Hampton Research	HR2-086	PEGRx HT	96

Znalezione warunki krystalizacyjne optymalizowałam w układzie kropli wiszącej nastawiając ręcznie znacznie większe krople krystalizacyjne (rzędu 1 - 8 μL). Optymalizacje polegały na nieznacznych zmianach warunków i dotyczyły temperatury krystalizacji, proporcji składników kropli (preparatu białka i roztworu rezerwuarowego), stężeń składników kropli (białka, pojedynczych odczynników lub jednocześnie całego roztworu) oraz pH.

Wykonywałam także szczepienia, przenosząc mikrokryształy, lub niewidoczne gołym okiem zarodki krystalizacji, do kropli tak przygotowanych, aby zamiast pojawiania się nowych kryształów (czyli nukleacji) następował jedynie wzrost przeniesionych obiektów. W trakcie szczepienia stosowałam różne techniki wykonania tego zabiegu. Przenosząc zarodki posługiwałam się włosem lub kocim wąsem. Szczepiłam krople ustabilizowane w układzie zamkniętym jak i świeżo przygotowane.

Optymalizacje krystalizacji miały na celu poprawić jakość uzyskiwanych kryształów, zmienić ich pokrój lub/i spowodować wzrost większych, łatwiejszych do wyłowienia, pojedynczych kryształów zamiast wysypu mikrokryształów.

4.5. Pomiary dyfrakcyjne

Wykorzystywanym przeze mnie źródłem promieniowania rentgenowskiego było promieniowanie synchrotronowe, które w biokrystalografii jest obecnie powszechnie stosowane. Charakteryzuje się ono wysoką intensywnością, która jest niezbędna w badaniach kryształów białek, ponieważ osiągają one niewielkie rozmiary przy jednocześnie stosunkowo dużych rozmiarach komórki elementarnej. Ponadto wielką zaletą promieniowania synchrotronowego jest możliwość dobrania odpowiedniej długości fali, co ma ogromne znaczenie przy pomiarach metodą dostrojonej dyfrakcji anomalnej (MAD lub SAD).

Podczas pomiarów kryształy były witryfikowane w strumieniu par azotu, co redukuje drgania termiczne atomów i ogranicza uszkodzenia radiacyjne wywołane dyfuzją wolnych rodników w sieci krystalicznej (**Garman, 1999**). Aby uniknąć powstawania lodu, przed zamrożeniem kryształ najczęściej nasącza się roztworem krioprotektanta, który dzięki dużej lepkości zapobiega krystalizacji wody.

W trakcie pomiaru, podczas gdy kryształ obracany jest wokół osi prostopadłej do wiązki rentgenowskiej, rejestruje się zwykle kilkaset obrazów dyfrakcyjnych. Rejestracja każdego pojedynczego obrazu wiąże się z niewielkim obrotem (0.2 - 2.0 º) i naświetlaniem kryształu przez określony czas (od ułamka do kilkudziesięciu sekund) w zależności od intensywności wiązki i zdolności dyfrakcyjnych kryształu. Aby zarejestrować kompletny zbiór danych, należy zmierzyć refleksy z odpowiedniego zakresu kąta, który zależy od symetrii grupy przestrzennej i orientacji kryształu. W celu zminimalizowania błędów pomiarowych zaleca się zarejestrowanie każdego refleksu więcej niż jeden raz poprzez zmierzenie intensywności refleksów symetrycznych.

4.6. Strategie rozwiązywania struktur krystalograficznych

Dyfrakcja promieni rentgena na kryształach jest efektem oddziaływania fali elektromagnetycznej z elektronami obecnymi w, wysoko uporządkowanym w sieci krystalicznej, zbiorze cząsteczek białka. Z równania Bragga wiadomo, że odbicie/ugięcie promieni o ustalonej długości fali (λ), przez rodziny płaszczyzn sieciowych (hkl) charakteryzowane przez odstępy międzypłaszczyznowe (d_{hkl}), zachodzi tyko dla ściśle określonego kąta odbłysku θ:

$$n\lambda = 2\, d_{hkl}\, \sin\theta$$

Obrazy otrzymane w wyniku eksperymentu dyfrakcyjnego zawierają refleksy, których geometryczny rozkład wynika z symetrii sieci krystalicznej, natomiast intensywności kodują informację o położeniu atomów białka w krysztale. Z matematycznego punktu widzenia, rozkład gęstości elektronowej $\rho(xyz)$ w komórce elementarnej (o objętości V) kryształu jest transformatą Fouriera czynników struktury $F(hkl)$, tj. wielkości zespolonych, określonych przez dwie składowe: amplitudy $|F(hkl)|$ i fazy $\varphi(hkl)$ fali rozproszonej:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)| \exp\, [-2\pi i(hx + ky + lz) + i\varphi(hkl)]$$

Czynnik struktury można wyznaczyć jako sumę Fouriera czynników rozpraszania atomowego (f_j) wszystkich N atomów (j) o współrzędnych x_j, y_j, z_j, w komórce elementarnej. Czynnik rozpraszania atomowego f_j zależy przede wszystkim od liczby elektronów atomu oraz od kąta rozpraszania θ.

$$F(hkl) = \sum_{j=1}^{N} fj\, \exp\, \left[2\pi i\left(hx_j + ky_j + lz_j\right)\right]$$

Intensywność refleksu $I(hkl)$ w obrazie dyfrakcyjnym stanowi odzwierciedlenie natężenia wiązki ugiętej i jest proporcjonalna do kwadratu amplitudy fali ugiętej. Zatem aby uzyskać informacje o amplitudach czynników struktury należy w odpowiednich programach komputerowych przeliczyć zgromadzone dane (przetworzyć obrazy dyfrakcji, Rozdział 4.6.1) wg. schematu:

$$I(hkl) \rightarrow\ |F(hkl)|$$

Niestety pomiar dyfrakcji promieniowania rentgenowskiego na kryształach białek nie pozwala na wyznaczenie eksperymentalnie kąta fazowego $\varphi(hkl)$. Jest to tak zwany problem fazowy, którego rozwiązanie wymaga zastosowania jednej z trzech metod podstawowych (Rozdziały 4.6.2, 4.6.3 i 4.6.4) lub ich kombinacji.

Każda z tych metod w mniejszym lub większym stopniu bazuje na tzw. mapie Pattersona, którą oblicza się wyłącznie na podstawie kwadratów amplitud czynników struktury, z pominięciem faz. Mapa Pattersona zawiera maksima/piki reprezentujące wektory międzyatomowe, a wysokości tych pików są proporcjonalne do iloczynu liczby elektronów wchodzących w skład atomów tworzących wektor międzyatomowy.

4.6.1. Przetwarzanie danych dyfrakcyjnych

Podczas przetwarzania obrazów dyfrakcyjnych, ustala się parametry komórki elementarnej i grupę przestrzenną, oraz tworzy listę refleksów niezależnych określonych przez wskaźniki *h, k, l* i intensywności (I) wraz z ich odchyleniami standardowymi (σI). W tym celu stosuje się programy takie jak *Mosflm* (**Leslie i Powell, 2007**) oraz *HKL* (**Otwinowski i Minor, 1997**), a także wykorzystywany przeze mnie *XDS* (**Kabsch, 2010**). Najważniejsze etapy procesu przetwarzania danych dyfrakcyjnych to:

- wskaźnikowanie, czyli nadanie refleksom wskaźników *h, k, l*, które określają „skąd" pochodzi informacja niesiona przez refleks;
- integracja, czyli odczytanie intensywności każdego refleksu, z uwzględnieniem intensywności pochodzącej od tła;
- skalowanie i uśrednianie, czyli sprowadzenie intensywności refleksów ze wszystkich obrazów dyfrakcyjnych do tej samej skali, dzięki czemu niweluje się wpływ np. wahań intensywności wiązki pierwotnej podczas pomiaru, czy uszkodzeń radiacyjnych.

Następnie zbiór intensywności refleksów (I) posłużył mi do stworzenia zbioru amplitud czynników struktury ($|F|$). Obliczenia wykonywałam w programie *XDSCONV*. Na tym etapie tworzyłam również zbiór refleksów wyłączonych z procesu udokładniania struktury służący do obliczania wskaźnika R_{free}, (**Kleywegt i Jones, 1997**) (Rozdział 4.7).

Parametry statystyczne uzyskiwane podczas przetwarzania danych (zebrane w Tabeli 4.3) dostarczają cennych informacji na temat jakości zbioru danych.

Tabela 4.3. Parametry statystyczne służące weryfikacji jakości zbioru danych dyfrakcyjnych

R_{merge}	- to parametr obliczany za pomocą wzoru: $\Sigma_{hkl} \Sigma_i \lvert I_i(hkl) - \langle I(hkl) \rangle \rvert \ / \ \Sigma_{hkl} \Sigma_i I_i(hkl)$ - mówi o tym, jaka jest różnica pomiędzy intensywnościami symetrycznych refleksów, które powinny być identyczne - wartość tego parametru powinna być jak najniższa
$\langle I/\sigma I \rangle$	- to średnia wartość sygnału do szumu - refleksy z obszaru rozdzielczości, w którym $\langle I/\sigma I \rangle$ spada poniżej 2 uznawane są za niewiarygodne
mozaikowość (º)	- średnie odchylenie od idealnego ułożenia krystalitów w krysztale
kompletność (%)	- liczba refleksów zarejestrowanych w stosunku do teoretycznej/spodziewanej liczby refleksów przy danej rozdzielczości - sumaryczna kompletność danych powinna wynosić co najmniej 85%
krotność = redundancja	- liczba refleksów zarejestrowanych w stosunku do liczby refleksów niezależnych, - liczba obserwacji danego refleksu niezależnego

4.6.2. Metoda podstawienia cząsteczkowego (MR)

Najprostszym i jedynym stosowanym przeze mnie sposobem rozwiązania problemu fazowego jest metoda podstawienia cząsteczkowego (MR, *ang. Molecular Replacement*). W metodzie tej strukturę białka badanego rozwiązuje się posiłkując się modelem tj. znaną strukturą białka podobnego do białka badanego. Identyczność sekwencji białek wynosząca minimum 25% daje realną nadzieję na rozwiązanie, ale go nie gwarantuje. Główną wadą metody MR jest to, że nie może ona zostać wykorzystana podczas badań nowych, nieznanych rodzin białek oraz to, że wprowadza trudne do wyeliminowania „skażenie" modelu.

Przykładowy algorytm metody MR obejmuje cztery etapy: (I) policzenie mapy Pattersona dla nieznanej struktury, (II) wygenerowanie mapy Pattersona dla znanej struktury modelowej, (III) wykonanie ich superpozycji, zmieniając orientację (rotację) i położenie (translację) modelu względem początku układu w nieznanej strukturze, (IV) przypisanie fazy odpowiednim refleksom. Umożliwia to obliczenie mapy gęstości elektronowej, na podstawie której buduje się model nieznanego białka.

Podstawienie cząsteczkowe wykonywałam korzystając z programu *PHASER* (**McCoy *et al.*, 2007**), który używa algorytmów opartych na metodzie największej wiarogodności (*ang. Maximum Likelihood*).

4.6.3. Metoda podstawienia izomorficznego (SIR i MIR)

Gdy brakuje odpowiedniej struktury modelowej, problem fazowy rozwiązuje się metodami opartymi na efektach rozpraszania przez atomy specjalne. Pierwszą z nich

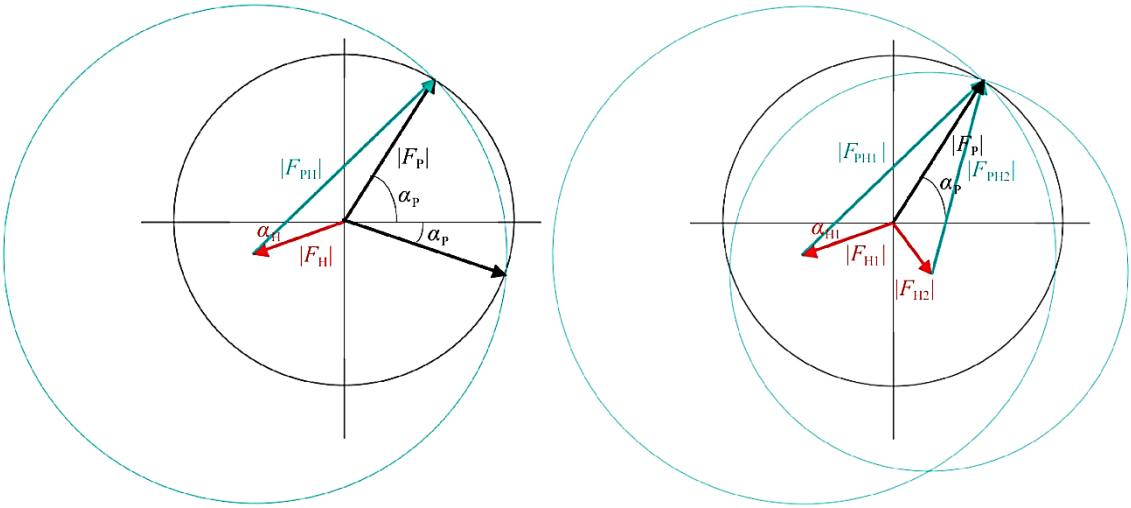
jest metoda podstawienia izomorficznego (SIR, *ang. Single Isomorphus Replacement*).

W metodzie SIR niezbędnym elementem jest pochodna izomorficzna, czyli kryształ, w którym poza cząsteczkami białka obecne są specyficznie przyłączone atomy ciężkie (np. Pt, Au lub Hg), zastępujące cząsteczki wody w otoczce hydratacyjnej. Kryształ ten musi jednocześnie posiadać identyczną postać jak kryształ natywny (identyczna grupa przestrzenna, upakowanie cząsteczek białka, bardzo podobne parametry komórki elementarnej). Pochodną izomorficzną można otrzymać poprzez współkrystalizację białka i jonów metalu ciężkiego lub też poprzez nasączanie kryształu natywnego odpowiednim roztworem. Następnie kryształ ten służy do rejestracji obrazów dyfrakcyjnych (F_{PH}) analogicznych do obrazów dyfrakcyjnych dla kryształu natywnego (F_P) prowadzonej w oparciu o tę samą teorię odbicia braggowskiego.

Pierwszym etapem w metodzie MIR jest wyznaczenie położeń atomów ciężkich na podstawie zbioru danych dla pochodnej izomorficznej. Przy wyznaczaniu tych położeń wykorzystuje się mapę Pattersona. Powstaje podstruktura składająca się wyłącznie z atomów specjalnych. Następnie znajomość położeń atomów ciężkich pozwala na wykorzystanie transformacji Fouriera do wyliczenia czynników struktury (F_H) dla tej podstruktury, a więc poznane zostają zarówno amplitudy jak i fazy dla podstruktury.

Następnie zdobyte informacje wykorzystuje się do wyznaczenia faz dla struktury białka. Ten etap najłatwiej zilustrować za pomocą konstrukcji Harkera (Rysunek 4.4), w której dyfrakcję od kryształu natywnego ($|F_P|$) i od pochodnej izomorficznej ($|F_{PH}|$) przedstawia się za pomocą przecinających się okręgów, których środki są przesunięte względem siebie o wektor $|F_H|$ reprezentujący różnicę w rozpraszaniu pochodnej i kryształu natywnego. W konstrukcji tej chodzi o znalezienie nieznanego kąta fazowego dla czynników struktury z kryształu natywnego (α_P), pod którym należy wykreślić wektor (F_P). Rozwiązania równania wektorowego $F_P + F_H = F_{PH}$ są dwa, dlatego konieczne jest zastosowanie różnych pochodnych izomorficznych (metoda MIR, *ang. Multiple Isomorphus Replacement*). Kolejna podstruktura zbudowana z atomów ciężkich innej pochodnej oznacza możliwość dorysowania do konstrukcji Harkera kolejnego wektora (F_{H2}) niosącego informację o jej atomach ciężkich i kolejnego okręgu ponownie przecinającego wcześniejsze okręgi w dwóch miejscach. Jednak wszystkie trzy okręgi będą się przecinały tylko w jednym miejscu, wyznaczając jednoznacznie kąt fazowy (α_P).

Rysunek 4.4. Konstrukcja Harkera obrazująca dwa rozwiązania w metodzie SIR i jedno w metodzie MIR. Rysunki pochodzą z publikacji w Acta Crystallographica (**Taylor, 2010)**.



4.6.4. Metoda dostrojonej dyfrakcji anomalnej (SAD i MAD)

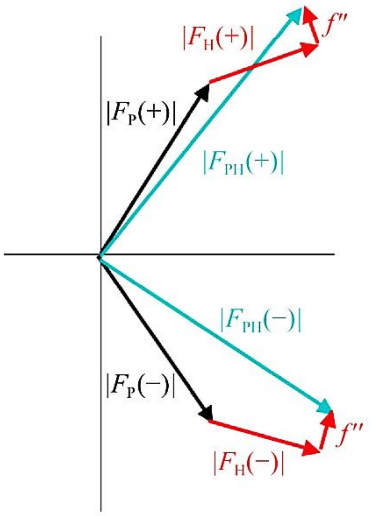
Innym sposobem rozwiązania problemu fazowego za pomocą atomów specjalnych jest wykorzystanie rozpraszania anomalnego (metoda SAD, *ang. Single-wavelength Anomalous Diffraction*).

Rozpraszanie anomalne ma miejsce wtedy, gdy dochodzi do rezonansu pomiędzy promieniowaniem rentgenowskim a naświetlanym obiektem. Promieniowanie jest wtedy silnie absorbowane przez jeden ze składników kryształu, czyli jednego rodzaju atomy specjalne, które zostają wzbudzone. Wzbudzenie to powoduje, że przestaje obowiązywać prawo Friedla, mówiące o jednakowych intensywnościach refleksów $I(hkl)$ i $I(-h-k-l)$, czyli tzw. par Friedla. Zatem analiza różnic intensywności $I(hkl)$ i $I(-h-k-l)$ na obrazach dyfrakcyjnych pozwala wyliczyć położenia atomów specjalnych rozpraszających anomalnie. Powstała podstruktura składająca się wyłącznie z „atomów anomalnych", podobnie jak w metodzie SIR, wykorzystywana jest do obliczeń, które ostatecznie prowadzą do wyznaczenia faz dla struktury białka (**Hendrickson i Ogata, 1997; Dauter *et al.*, 2002**).

W sytuacji anomalnej, różnice pomiędzy parami Friedla [$(F_{PH})^+$ i $(F_{PH})^-$] są wynikiem sumarycznego efektu normalnego rozpraszania zwykłych atomów (F_P) i rozpraszania wzbudzonych „atomów anomalnych" (F_H) (Rysunek 4.5), których czynniki rozpraszania atomowego składają się z trzech elementów:

- f_o - części rzeczywistej rozpraszania atomowego, która jest zależna tylko od kąta odbłysku (θ), lecz nie od długości fali,

- *f'* - części rzeczywistej poprawki anomalnej, która jest zależna od długości fali (λ) i której przesunięcie w fazie o 180 º względem f_o na wykresie Arganda powoduje skrócenie wektora f_o, stąd wspominane części rozpraszania (f_o - *f'*) prezentowane są jako jeden wektor,
- *f''* - części urojonej poprawki anomalnej, która jest zależna od długości fali (λ) i której przesunięcie fazowe względem f_o wynosi 90 º.

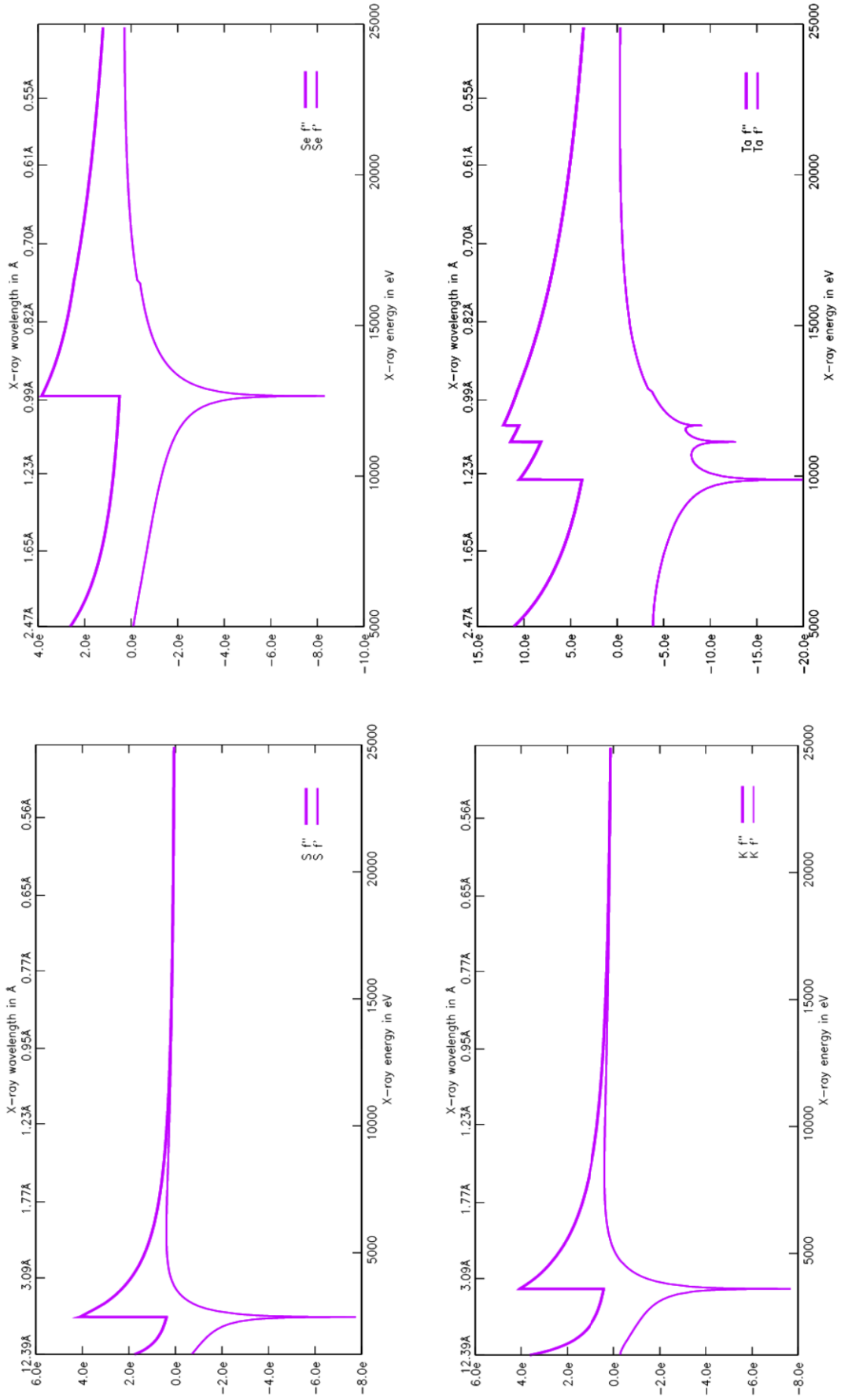


Rysunek 4.5. Konstrukcja Arganda dla refleksów pary friedelowskiej uwzględniająca rozpraszanie anomalne. Rysunek pochodzi z publikacji w Acta Crystallographica (**Taylor, 2010)**.

Zaobserwowanie zjawiska anomalnego jest najłatwiejsze podczas naświetlania kryształu promieniowaniem o długości fali (λ) odpowiadającej maksimum absorpcji atomu anomalnego występującego w krysztale. W wyborze λ pomagają wykresy progów absorpcji wyznaczone dla różnych pierwiastków, na których konkretnej długości fali odpowiadają zestawy poprawek (*f'* + *f''*) (Rysunek 4.6). Wartości poprawek anomalnych są znikome dla typowych długości fali i dla najczęściej występujących w białkach atomów pierwiastków lekkich (C, N, O), jednak już od atomów siarki są one wystarczająco wyraźne aby możliwe było ich użycie do rozwiązania problemu fazowego.

W metodzie SAD wykorzystuje się różne „anomalne atomy specjalne", które wprowadzane są do kryształów różnymi metodami. Kompleks bromku tantalu $[Ta_6Br_{12}]Br_2$ umieszcza się w krysztale, podobnie jak w metodzie SIR, nasączając go odpowiednim roztworem. Selen wprowadzany jest do struktury metodami biotechnologicznymi, już na etapie ekspresji białka, podczas której istniejące w sekwencji reszty metioniny są zastępowane przez Se-Met. Wykorzystuje się również natywne kryształy białka, w których wzbudza się atomy siarki, naturalnie występujące w resztach Cys i Met. Próg absorpcji atomów siarki odpowiada energii promieniowania rentgenowskiego o długości fali λ = 5.02 Å (Rysunek 4.6), niedostępnej w ramach pomiarów dyfrakcyjnych, z tego względu wykorzystuje się efekty anomalne przy λ z zakresu 1.9 - 2.1 Å. Sygnał anomalny atomów siarki jest stosunkowo słaby i wymaga pomiaru danych dyfrakcyjnych o najwyższej jakości, mierzonych z dużą redundancją.

Rysunek 4.6. Krzywe absorpcji dla atomów siarki, potasu, selenu i tantalu, wygenerowane przez Edgeplots web tool http://skuld.bmsc.washington.edu/scatter/.



W metodzie SAD, tak jak w metodzie SIR, otrzymuje się dwuznaczne rozwiązanie problemu fazowego. Tradycyjnym sposobem uzyskania jednoznaczności faz jest metoda MAD (*ang. Multi-wavelength Anomalous Diffraction*), dla której rejestruje się kilka zbiorów danych anomalnych naświetlając jeden kryształ promieniowaniem rentgenowskim o różnych długościach fali. Z kolei metoda SIRAS (*ang. Single Isomorphous Replacement with Anomalous Scattering*) łączy wykorzystanie efektów anomalnych z izomorficznym rozpraszaniem atomu ciężkiego. Obecnie, bardzo dokładne pomiary różnic anomalnych oraz coraz lepsze narzędzia obliczeniowe pozwalają wyjaśnić niejednoznaczność rozwiązania z metody SAD dzięki zastosowaniu procedury modyfikacji mapy gęstości elektronowej (**Taylor, 2010)**.

Efekty anomalne można wykorzystać nie tylko do rozwiązywana problemu fazowego, ale także do weryfikacji położeń atomów specjalnych w udokładnionym już modelu struktury białka. W tym celu jako współczynniki mapy gęstości elektronowej stosuje się różnice anomalne. Do uzyskania „anomalnych gęstości elektronowych" wystarcza nawet bardzo słaby sygnał anomalny. W swoich doświadczeniach do lokalizacji atomów specjalnych zastosowałam program *ANODE* (**Thorn i Sheldrick, 2011**).

4.7. Udokładnianie modelu struktury

Rozwiązanie problemu fazowego umożliwia wyliczenie rozkładu gęstości elektronowej w krysztale i stworzenie wstępnego modelu atomowego struktury białka, który następnie podlega procesowi udokładniania. Celem udokładniania jest takie zmodyfikowanie parametrów modelu białka, aby jak najwierniej odzwierciedlał on dane eksperymentalne. Procedura obejmuje dwa cyklicznie powtarzane etapy. Pierwszym jest ręczna przebudowa modelu w oparciu o mapy gęstości elektronowej o współczynnikach $2F_o$ - F_c i F_o - F_c, którą wykonywałam w programie *COOT* (**Emsley i Cowtan, 2004**). Drugim etapem są automatyczne obliczenia minimalizujące rozbieżność między obliczonymi i zmierzonymi amplitudami czynników struktury. Obliczenia te wykonywałam w programach *PHENIX* (**Adams et al., 2010**) oraz *REFMAC5* (**Murshudov et al., 2011**) z pakietu *CCP4* (**Winn et al., 2011**). W opcji „największej wiarogodności" programy te poszukują takiego modelu, który maksymalizuje prawdopodobieństwo zarejestrowania takiego zestawu danych, jaki właśnie został zarejestrowany.

Wprowadzane w udokładnianym modelu zmiany dotyczyły głównie reszt aminokwasowych. Poza tym dobudowywałam również związki niskocząsteczkowe

(cząsteczki wody, kationy, aniony, ligandy, itp.) w oparciu o wielkość i kształt zapełnianych gęstości elektronowych, geometrię potencjalnych oddziaływań oraz wiedzę o użytych podczas oczyszczania i krystalizacji odczynnikach.

Sprawdzając tożsamość wprowadzanych do struktury jonów metali korzystałam między innymi z dostępnego przez Internet serwera CMM (*ang. CheckMyMetal*, **Zheng et al., 2014**). Identyfikacje jonów ułatwiała mi także metoda BVM (*ang. Bond Valence Method*, **Brese i O'Keeffe, 1991**) polegająca na weryfikacji dwóch parametrów. Parametry te obliczałam na podstawie długości poszczególnych wiązań koordynacyjnych, a wyniki porównywałam ze znanymi wartościami idealnymi dla tych parametrów. Pierwszym parametrem jest walencyjność wiązań (R_{ij}), zależna od tożsamości jonu metalu (i) i tożsamości (jednego) rodzaju atomów z którymi oddziałuje (j). Rekomendowane wartości R_{ij} zostały określone i stabelaryzowane (**Brese i O'Keeffe, 1991**). Drugi parametr to walencyjność kationu metalu (V_i), która powinna być tożsama z jego ładunkiem.

$$R_{ij} = b \ln\left[\frac{V_{iR}}{\sum_j \exp\left(\frac{-d_{ij}}{b}\right)}\right] \qquad V_i = \sum_j \exp\left[\frac{R_{jiR} - d_{ij}}{b}\right]$$

R_{jiR} - rekomendowana wartość walencyjności wiązań pomiędzy atomami i i j

V_{iR} - rekomendowana walencyjność atomu i

d_{ij} - obserwowana długość wiązania między atomami i i j

b - stała uniwersalna wynosząca 0.37 Å.

Ze względu na dużą liczbę atomów w strukturze białka, liczba udokładnianych parametrów również jest duża. Dlatego do procesu udokładniania włącza się dodatkowe informacje. Prawie zawsze konieczne jest zastosowanie więzów, uwzględniających wcześniej zdefiniowane powszechnie dostępne informacje dotyczące stereochemii składników białek (długości wiązań, wartości kątów walencyjnych, itp.). O tym, jak silne więzy należy zastosować decyduje rozdzielczość zbioru danych. Oprócz więzów podczas badań korzystałam także z parametrów TLS (*ang. translation, libration, screw motion*), które zakładają, że grupa atomów może poruszać się jak bryła sztywna. Osobne grupy TLS mogą stanowić zarówno całe podjednostki białek, domeny lub ich fragmenty. Parametry TLS częściowo odzwierciedlają anizotropię skorelowanych drgań całych grup atomów.

4.8. Walidacja i ocena jakości modelu

Podstawowym kryterium określającym zgodność pomiędzy obliczonymi (na podstawie modelu) i obserwowanymi czynnikami struktury są krystalograficzne wskaźniki rozbieżności R i R_{free} (**Brünger, 1992**):

$$R = \frac{\sum_{(hkl)} w_{(hkl)} \left| |F_o(hkl)| - k|F_c(hkl)| \right|}{\sum_{(hkl)} |F_o(hkl)|}$$

$$R_{free} = \frac{\sum_{(hkl) \in T} w_{(hkl)} \left| |F_o(hkl)| - k|F_c(hkl)| \right|}{\sum_{(hkl) \in T} |F_o(hkl)|}$$

R - krystalograficzny liniowy wskaźnik rozbieżności

R_{free} - testowy wskaźnik rozbieżności

$w_{(hkl)}$ - waga przypisana obserwacji *hkl*

$|F_o(hkl)|$ - moduł obserwowanego czynnika struktury

$|F_c(hkl)|$ - moduł obliczonego czynnika struktury

k - czynnik skali

hkl - wskaźniki *hkl* refleksu

T - zbiór wskaźników refleksów testowych

Obserwacja wskaźników rozbieżności pozwala ocenić, czy przebudowa modelu zmierza w dobrym kierunku. Wskazuje również czy wprowadzenie dodatkowych parametrów optymalizowanych w udokładnianiu jest usprawiedliwione dostateczną liczbą obserwacji i czy nie dochodzi do nadinterpretacji, o której świadczy wzrost wartości R_{free} towarzyszący obniżeniu R. Przyjmuje się, że pod koniec udokładniania wskaźniki te powinny być jak najniższe, choć ich wielkość jest mocno uzależniona od rozdzielczości i jakości danych eksperymentalnych, natomiast różnica pomiędzy nimi nie powinna przekraczać 8%. Podczas obliczania wskaźnika R_{free} korzysta się z pewnej puli refleksów testowych wykluczonych z procesu udokładniania, które zostały wybrane losowo z całego zbioru danych dyfrakcyjnych już na etapie procesowania.

Kolejnym kryterium określającym jakość modelu jest średnie kwadratowe odchylenie (r.m.s.d.) poszczególnych parametrów, takich jak długość wiązań i kąty walencyjne, od wartości idealnych. Ponadto analiza wykresu Ramachandrana, który jest graficzną reprezentacją kątów torsyjnych φ/ψ łańcucha głównego, pozwala na szybką

identyfikację reszt o nieprawidłowej konformacji.

Do weryfikacji stereochemicznej jakości modelu zastosowałam programy Procheck (**Laskowski et al., 1993**) i MolProbity (**Chen et al., 2010**), które są dostępne na serwerze ADIT.

4.9. Analiza struktur i prezentacja wyników

Przeprowadzając analizę struktur oraz przygotowując rysunki i schematy opisujące wyciągnięte z tej analizy wnioski, posługiwałam się rozmaitymi programami podanymi w Tabeli 4.4.

Tabela 4.4. Programy i serwery, które wykorzystywałam do analizy i opisu struktur.

program / serwer	przeznaczenie
SSM (**Krissinel i Henrick, 2004**)	porównywanie struktur (w programie *COOT*)
BLAST (**Altschul *et al.,* 1997**)	wyszukiwanie i porównywanie sekwencji aminokwasowych
*ClustalO (**McWilliam et al., 2013***	porównywanie sekwencji aminokwasowych
ALIGN (**Satow *et al.,* 1986**)	nakładanie struktur metodą najmniejszych kwadratów
DSSP (**Kabsh i Sander, 1983**)	analiza struktury drugorzędowej
PDBsum (**Laskowski, 2009**)	analiza struktury czwartorzędowej
PyMOL (**DeLano, 2002**)	przygotowanie ilustracji

5. Bakteryjne białka nodulacji

5.1. Pięć obiektów badań w kilkunastu wersjach

Celem projektu, który dotyczył bakteryjnych białek nodulacji, było przetestowanie tych białek pod kątem ich przydatności do badań krystalograficznych. Nad projektem pracowało w sumie kilka osób, które zajmowały się łącznie kilkunastoma białkami nodulacji. Moim zadaniem było przetestowanie pięciu enzymów: NodA, NodB, NodU, NolO i NoeI.

Zakres moich działań obejmował prace od etapu pozyskania materiału genetycznego dla wybranego enzymu, aż do etapu krystalizacji. Do badań wykorzystałam kilka szczepów symbiotycznych bakterii brodawkowych, w których homologi interesujących mnie białek miały nieco odmienne sekwencje. Ponadto stosowane wektory pozwalały na uzyskanie tych samych białek z różnymi cechami dodatkowymi, np. ze znacznikiem powinowactwa lub białkiem fuzyjnym. Zatem każdy z docelowych pięciu enzymów został przeze mnie przetestowany w kilku różnych wersjach.

Prace z różnymi wersjami danego białka rozpoczynałam w różnym czasie, a postęp tych prac miał różne tempo. Nowe warianty tworzyłam dopiero wtedy, gdy poprzednie zostały już przetestowane. Stąd wykonywane przeze mnie prace były prowadzone na wielu etapach jednocześnie przez cały okres trwania projektu. Opisane w części doświadczalnej zadania są jednak przedstawione w sposób zbiorczy, etapami. Zebrane są tematycznie i niekoniecznie odzwierciedlają chronologię prac.

5.2. Część doświadczalna

5.2.1. Pozyskanie materiału genetycznego rizobiów

Cztery szczepy symbiotycznych bakterii glebowych, z których pozyskiwałam geny dla wybranych białek nodulacji (*Sinorhizobium meliloti* 1021, *Bradyrhizobium* sp. WM9, *Rhizobium* sp. NGR234 i *Bradyrhizobium japonicum* USDA110), otrzymałam od prof. Tomasza Stępkowskiego. Bakterie były w postaci próbek płynnych hodowli komórkowych zamrożonych w -80ºC z 15% dodatkiem glicerolu. Wszystkie cztery

szczepy były przeze mnie hodowane w temperaturze 27ºC, na pożywce YMB, początkowo stałej (z agarem), a następnie płynnej (z wytrząsaniem 300 rpm), przy czym czas hodowli wynosił, w zależności od szczepu, od dwóch do siedmiu dni. Komórki z hodowli płynnych posłużyły mi do izolacji całkowitego DNA bakteryjnego (Tabela 5.1) oraz posiewu testowego na pożywce stałej, dzięki któremu wykluczyłam zakażenia.

Materiał genetyczny wyizolowany z czterech szczepów rizobiów wykorzystywałam jako matrycę w reakcjach PCR, za pomocą których tworzyłam inserty DNA z genami wybranych białek nodulacji, przeznaczone do wklonowania do wektorów plazmidowych.

Tabela 5.1. Protokół izolacji całkowitego DNA bakteryjnego.

do próbówki eppendorf (2 mL) przenieść 1 mL płynnej hodowli bakteryjnej, zwirować hodowlę (5 min., 4ºC, 13 000 rpm) i usunąć supernatant
osad komórkowy zawiesić w 1 mL buforu STE (10 mM TrisHCl, 1 mM EDTA, 0.5 mM NaCl, pH 7), zwirować zawiesinę (5 min., 4ºC, 13 000 rpm) i usunąć supernatant
osad ponownie zawiesić w 1 mL buforu STE, inkubować 10 min na lodzie, zwirować zawiesinę (5 min., 4ºC, 13 000 rpm) i usunąć supernatant
osad zawiesić w 450 µL buforu TE (10 mM TrisHCl, 1 mM EDTA, pH 8), dodać 50 µL preparatu lizozymu w buforze TE (o stężeniu 4 mg/mL, czyli 0.2 mg enzymu) i inkubować 1 godzinę w 37ºC wytrząsając 300 rpm
dodać 10 µL preparatu proteinazy K i 110 µL 10% roztworu SDS inkubować 1-3 godzin w 37ºC wytrząsając 300 rpm, do czasu aż mieszanina stanie się lepka i klarowna
dodać 110 µL 5 M NaCl i intensywnie wytrząsać (vortex), dodać 90 µL CTAB/NaCl i intensywnie wytrząsać (vortex), inkubować przez 10 min. w 65ºC
dodać 820 µL chloroformu, intensywnie wytrząsać (vortex) i zwirować (5 min., 4ºC, 13 000 rpm), zebrać fazę wodną do nowych próbówek eppendorf (2 mL)
dodać 800 µL chloroformu, intensywnie wytrząsać (vortex) i zwirować (5 min., 4ºC, 13 000 rpm), zebrać fazę wodną do nowych próbówek eppendorf (1 mL)
dodać 570 µL izopropanolu, zmieszać, inkubować 10-20 min. na lodzie, zwirować (10 min., 15ºC, 13 000 rpm) i bardzo delikatnie usunąć supernatant
przemyć osad z DNA 1 mL 80% etanolu, zwirować (5 min., 4ºC, 13 000 rpm) i bardzo delikatnie usunąć supernatant
dokładnie osuszyć osad z DNA
zawiesić osad z DNA w 50 µL wody i pozostawić preparat w 4ºC przez noc celem rozpuszczenia

5.2.2. Tworzenie wektorów ekspresyjnych

Wszystkie przygotowane przeze mnie plazmidowe wektory ekspresyjne umożliwiające ekspresję rekombinowanych białek nodulacji (Tabela 5.2) można zaklasyfikować do jednego z trzech typów wektorów: pET, pMAL, pMB, które różnią się cechami, jakie nadają eksprymowanym dzięki nim białkom. Poszczególne typy

wektorów przygotowywałam w nieco odmienny sposób, po czym stosowałam ten sam schemat działania umożliwiający wybór i weryfikację wektora o poprawnej konstrukcji.

Tabela 5.2. Zestawienie wektorów ekspresyjnych przygotowanych dla białek nodulacji.

nazwa białka nodulacji	organizm źródłowy i numer identyfikacyjny białka*	masa białka natyw-nego (kDa)	wykorzystany wektor plazmidowy (cecha dodana do białka)	masa białka rekombi-nowanego (kDa)	nazwa wektora ekspresyjnego zawierającego gen białka nodulacji
NodA	*Sinorhizobium meliloti* 1021 NP_435721.1	21.8	pET151/D-topo (znacznik histydynowy)	25.8	pET-mel-NodA
	Bradyrhizobium sp. WM9 AAK00155.1	23.4	pET151/D-topo (znacznik histydynowy)	27.4	pET-wm9-NodA
			pMALc5x (białko fuzyjne MBP)	65.9	pMAL-wm9-NodA
			pMB (białko fuzyjne MBP)	63.4	pMB-wm9-NodA
NodU	*Rhizobium* sp. NGR234 AAB91782.1	60.5	pET151/D-topo (znacznik histydynowy)	64.5	pET-ngr-NodU
	Bradyrhizobium sp. WM9 AAK00159.1	62.0	pET151/D-topo (znacznik histydynowy)	66.0	pET-wm9-NodU
			pMALc5x (białko fuzyjne MBP)	104.5	pMAL-wm9-NodU
NolO	*Rhizobium* sp. NGR234 AAB91692.1	75.5	pET151/D-topo (znacznik histydynowy)	79.5	pET-ngr-NolO
	Bradyrhizobium sp. WM9 AAK00162.1	74.8	pET151/D-topo (znacznik histydynowy)	78.8	pET-wm9-NolO
			pMALc5x (białko fuzyjne MBP)	117.3	pMAL-wm9-NolO
NodB	*Bradyrhizobium* sp. WM9 AAK00156.1	23.5	pET151/D-topo (znacznik histydynowy)	27.5	pET-wm9-NodB
			pMALc5x (białko fuzyjne MBP)	66.0	pMAL-wm9-NodB
			pMB (białko fuzyjne MBP)	63.5	pMB-wm9-NodB
NoeI	*Bradyrhizobium japonicum* USDA110 BAC47327.1	24.4	pET151/D-topo (znacznik histydynowy)	28.4	pET-jap-NoeI
			pMALc5x (białko fuzyjne MBP)	66.9	pMAL-jap-NoeI
			pMB (białko fuzyjne MBP)	64.4	pMB-jap-NoeI

* numer z internetowej bazy: centrum informacji biotechnologicznej NCBI (*ang. National Center for Biotechnology Information*)

5.2.2.1. Przygotowywanie wektorów ekspresyjnych typu pET

Wektory typu pET umożliwiają ekspresję docelowego białka z N-terminalnym znacznikiem histydynowym, który powiększa masę rekombinowanego białka o ok. 4 kDa. Znacznik ten może zostać usunięty za pomocą proteazy TEV, przy czym po cięciu

na aminowym końcu białka pozostaje sześć dodatkowych reszt aminokwasowych (GIDPFT-).

Tworząc wektory pET, w pierwszej kolejności przygotowywałam inserty DNA. Za pomocą reakcji PCR z materiału genetycznego rizobiów (np. *Bradyrhizobium* sp. WM9) wyodrębniałam fragment DNA stanowiący gen interesującego mnie białka nodulacji (np. białka NodA). Podczas tych reakcji wykorzystywałam polimerazę PWO (Roche) oraz startery (F- forward i R- reverse) zaprojektowane tak, aby pasowały do początku i końca sekwencji wyodrębnianych genów, przy czym starter F wprowadzał dodatkowe cztery nukleotydy CACC na 5' końcu nici kodującej insertów.

Próbki produktów reakcji PCR (insertów) poddawałam elektroforezie na 1% żelu agarozowym celem potwierdzenia otrzymania zamplifikowanych fragmentów DNA. Jeśli uzyskiwałam kilka produktów, całość mieszaniny poreakcyjnej rozdzielałam elektroforetycznie, a następnie izolowałam z żelu jedynie te fragmenty DNA, które miały odpowiednią długość. Inserty izolowane były za pomocą komercyjnego zestawu QIAquick Gel Extraction Kit firmy Quiagen oraz wirówki Eppendorf.

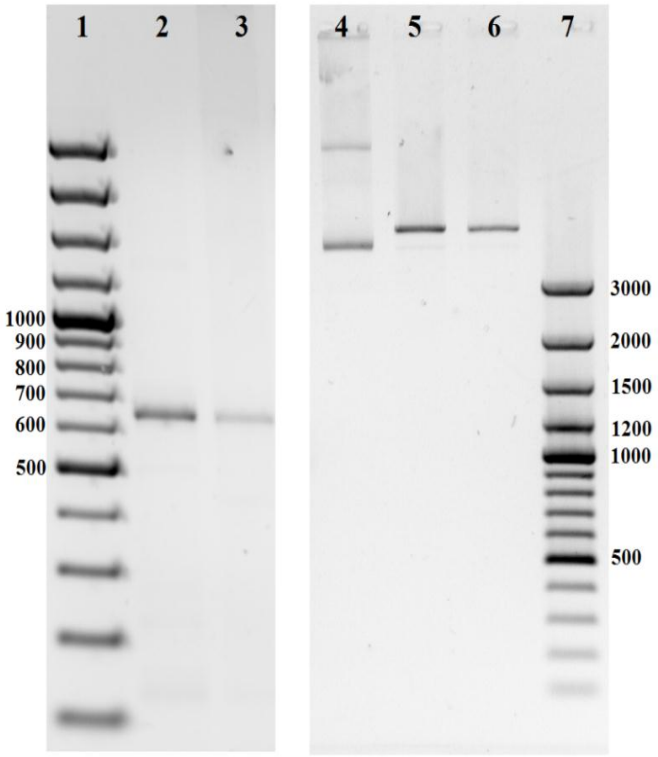
Ligacja insertu DNA i plazmidowego wektora pET151/D-topo (Invitrogen) następowała na skutek obecności topoizomerazy (Rozdział 4.1.2.2), a wprowadzone do insertu dodatkowe nukleotydy (5' CACC-) zapewniały poprawność kierunku ligacji dzięki komplementarności tylko do jednego z końców wektora. Reakcje ligacji, które przeprowadziłam zgodnie z protokołem dostarczonym wraz z wektorem pET151/D-topo przez firmę Invitrogen, zasadniczo polegały na zmieszaniu insertu z wektorem i trwały zaledwie 5 minut.

5.2.2.2. Przygotowywanie wektorów ekspresyjnych typu pMAL

Wektory typu pMAL umożliwiają ekspresję docelowego białka z N-terminalnym białkiem fuzyjnym MBP, przy czym oba białka oddzielone są dwudziesto-aminokwasowym łącznikiem, a masa całej konstrukcji jest większa od masy docelowego białka o ok. 42.5 kDa. Białko fuzyjne można odciąć używając czynnik Xa, nie pozostawiając żadnego dodatkowego aminokwasu w produkcie docelowym.

Podobnie jak poprzednio, przygotowywanie wektorów pMAL rozpoczęłam od reakcji PCR, w której matrycą był materiał genetyczny rizobiów, używanym enzymem polimeraza PWO (Roche), a startery (F i R) umożliwiały wyłowienie konkretnych genów. Zamplifikowane inserty posiadały, bezpośrednio po sekwencji genu, 12

dodatkowych par zasad, stanowiących miejsce cięcia enzymem SbfI, wprowadzone za pomocą startera R. Tak jak poprzednio, długość insertów weryfikowałam za pomocą elektroforezy (Rysunek 5.1).



Rysunek. 5.1. Przykładowe żele agarozowe weryfikujące postęp prac związanych z przygotowywaniem wektora ekspresyjnego typu pMAL. Żele dotyczą tworzenia pMAL-wm9-NodA, czyli wektora ekspresyjnego dla białka NodA z *Bradyrhizobium* sp. WM9. Ścieżki na żelu agarozowym: 1 i 7– marker DNA, dla najistotniejszych prążków podałam liczbę par zasad (p.z.), 2 i 3– NodA (gen ma 633 p.z.), 2- produkt reakcji PCR (645 p.z.), 3- insert po cięciu SbfI, 4, 5 i 6– wektor pMAL, 4- kolisty zamknięty plazmid (5677 p.z.) migrujący w postaci rozluźnionej (wyższy prążek) i superskręconej (niższy prążek) , 5- wektor po pierwszym cięciu enzymem XmnI (5677 p.z.) i 6- wektor po drugim cięciu enzymem SbfI (5622 p.z.). Wektor w postaci liniowej migruje inaczej niż zamknięty – prążek pojawia się na żelu pomiędzy prążkami dwóch form postaci kolistej.

Tabela 5.3. Oczyszczanie fragmentów DNA po reakcjach enzymatycznych. Procedura wykonywana w temperaturze pokojowej, chyba że zaznaczyłam inaczej.

do próbówki eppendorf (1.5 mL) przenieść mieszaninę reakcyjną, dodać wodę do objętości końcowej 100 µL
dodać 100 µL fenolu (pH 8.0), intensywnie wytrząsać 2 min. (vortex), zwirować (1 min., 10 000 rpm), zebrać fazę wodną do nowych próbówek eppendorf (1.5 mL)
dodać 100 µL chloroformu, intensywnie wytrząsać 2 min. (vortex), zwirować (1 min., 10 000 rpm), zebrać fazę wodną do nowych próbówek eppendorff (1.5 mL)
dodać 100 µL chloroformu, intensywnie wytrząsać 2 min. (vortex), zwirować (1 min., 10 000 rpm), zebrać fazę wodną do nowych próbówek eppendorff (1.5 mL)
dodać 10 µL 3 M octanu sodu (pH 5.2) i 300 µL 96% lodowatego etanolu, intensywnie wytrząsać 3 godz. w temperaturze 4ºC (vortex), zwirować (2 godz., 4ºC, 14 000 rpm) i bardzo delikatnie usunąć supernatant
dokładnie osuszyć osad z DNA
zawiesić osad z DNA w 10 µL wody wytrząsając 10 min., 4ºC (vortex)

Inserty DNA zawarte w mieszaninach reakcyjnych po reakcjach PCR oczyszczałam (Tabela 5.3), a następnie poddawałam cięciu enzymem restrykcyjnym SbfI (Fermentas) i ponownie oczyszczałam (Tabela 5.3). W ten sposób otrzymywałam fragmenty DNA posiadające jeden lepki koniec. Wektor pMALc5x (NEB) poddawałam

cięciu dwoma enzymami restrykcyjnymi: XmnI (Promega), który tworzył tępy koniec wektora oraz SbfI (Fermentas), który tworzył lepki koniec wektora, skracając jednocześnie wektor o 55 p.z., czyli o długość wyciętego fragmentu mcs (*ang. multiple cloning site*). Po reakcjach przeprowadzonych zgodnie z protokołami producentów enzymów, wektor oczyszczałam (Tabela 5.3), a następnie od razu przechodziłam do reakcji ligacji. Ligacja insertu DNA i wektora pMALc5x następowała w wyniku działania T4 DNA ligazy (Fermentas) i była przeprowadzona zgodnie z protokołem dostarczonym wraz z enzymem. Dwa różne końce (tępy i lepki) w insercie i wektorze zapewniały poprawność kierunku połączenia tych fragmentów DNA.

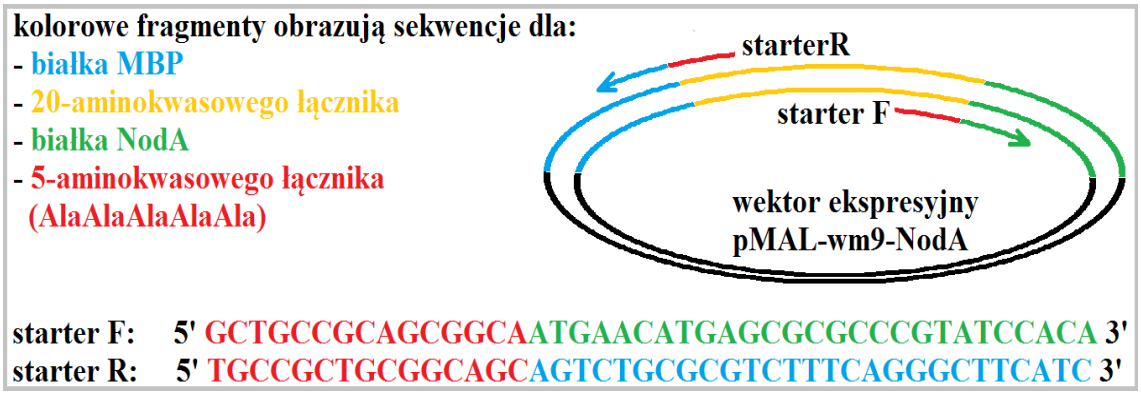
5.2.2.3. Przygotowywanie wektorów ekspresyjnych typu pMB

Wektory typu pMB umożliwiają ekspresję docelowego białka z N-terminalnym białkiem fuzyjnym MBP, przy czym oba białka oddzielone są fragmentem łańcucha polipeptydowego składającego się z zaledwie pięciu reszt alaniny, co powoduje usztywnienie i stabilizację całej konstrukcji. Brak jest miejsca cięcia proteazą, które pozwoliłoby odseparować oba białka. Masa produktu fuzyjnego powiększona jest o ok. 40 kDa, w stosunku do masy białka docelowego.

Podczas opracowywania sposobu przygotowania wektorów typu pMB opierałam się na regułach metody PIPE (Rozdział 4.1.3). Opracowany plan działania składał się zaledwie z jednego etapu – jednej reakcji PCR (Rysunek 5.2). Podczas tej reakcji matrycą dla polimerazy Pfu (Thermo Scientific) były wektory ekspresyjne typu pMAL z wklonowanymi już insertami zawierającymi sekwencje genów dla badanych białek nodulacji. Trwające prawie osiem godzin reakcje PCR polegały na amplifikacji niemal całych wektorów matrycowych. Specjalnie zaprojektowane startery (F i R) nie dopuszczały do amplifikacji tylko tego fragmentu wektora matrycowego, który kodował dwudziestoaminokwasowy łącznik pomiędzy białkiem MBP i docelowym białkiem nodulacji. Jednocześnie startery te w miejscu starego łącznika wprowadzały nowy piętnastonukleotydowy fragment DNA. Sekwencja tego nowego fragmentu była odpowiedzialna za obecność krótkiego i sztywnego pięcioalaninowego łącznika w rekombinowanych białkach ulegających ekspresji z wektorów typu pMB. Dodatkowo, piętnastonukleotydowe fragmenty starterów F i R zawierające sekwencję nowego łącznika były komplementarne i umożliwiały połączenie liniowego produktu reakcji PCR w kolistą cząsteczkę plazmidowego wektora ekspresyjnego typu pMB. Fragmenty

te pod koniec reakcji PCR pozostawały jednoniciowe dzięki pominięciu ostatniego etapu reakcji, czyli etapu uzupełniania matryc (*ang. final extension*).

Rysunek 5.2. Schemat obrazujący przebieg reakcji PCR podczas tworzenia wektora ekspresyjnego typu pMB na przykładzie pMB-wm9-NodA.



5.2.2.4. Weryfikacja i wybór wektorów o poprawnej konstrukcji i sekwencji

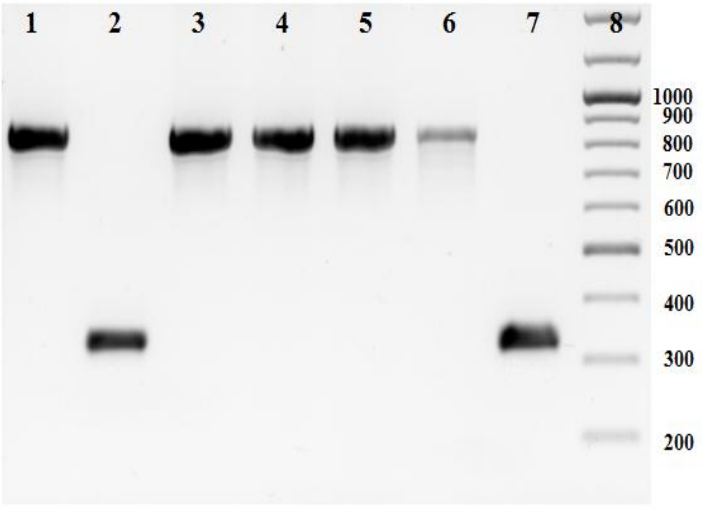
Proces przygotowywania plazmidowych wektorów ekspresyjnych zazwyczaj nie przebiega w stu procentach poprawnie. W trakcie kolejnych etapów pracy mogą pojawić się błędy w sekwencjach (np. podczas reakcji PCR) lub błędy w konstrukcji wektora (np. podczas cięcia enzymami restrykcyjnymi, ligacji). Zatem wszystkie przygotowane przeze mnie plazmidowe wektory ekspresyjne należące do trzech typów (pET, pMAL, pMB) podlegały takiemu samemu procesowi weryfikacji (Tabela 5.4).

Aby wyodrębnić poprawne wektory ekspresyjne z danej mieszaniny poreakcyjnej (po reakcji ligacji w przypadku wektorów pET i pMAL lub po reakcji PCR w przypadku wektorów pMB) wykonywałam transformację do komórek kompetentnych, zakładając, że jedna komórka przyjmie jeden wektor. Kolonie tworzone przez rozmnożone, zawierające identyczne wektory komórki, służyły następnie jako materiał do dwóch testów weryfikujących poprawność wektorów, które zawierały.

W pierwszym teście wykorzystywałam reakcję PCR do amplifikacji fragmentu wektora zawierającego insert oraz przylegające do niego od strony 5’ i 3’ końca dodatkowe sekwencje, obie o długości ok. 100 - 200 par zasad. Odpowiednia długość produktów reakcji PCR wskazywała wektory o poprawnej konstrukcji, które poddawałam kolejnemu testowi (Rysunek 5.3). W drugim teście za pomocą sekwencjonowania sprawdzałam poprawność sekwencji nukleotydowej tych samych fragmentów wektorów co w teście pierwszym.

Tabela 5.4. Protokół służący weryfikacji poprawności plazmidowych wektorów ekspresyjnych (pET, pMAL, pMB) zawierających inserty umożliwiające ekspresję bakteryjnych białek nodulacji.

etap	stosowane materiały	procedura
transformacja	mieszanina poreakcyjna zawierające plazmidowe wektory ekspresyjne komórki top10 (INVITROGEN) pożywka płynna SOC	- delikatne zmieszanie rozmrożonych komórek (50 µL) z wektorami (ok. 20 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5 min. - zmieszanie komórek z SOC (250 µL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla pojedynczych kolonii	pożywka stała LB z ampicyliną (0.1 mg/mL)	- delikatne rozprowadzenie hodowli 50 µL, 100 µL i 150 µL odpowiednio na trzech szalkach Petriego - inkubacja 37ºC, 16 godz. - wybór kilku pojedynczych kolonii, nadanie im numerów i przeszczepienie na świeżą szalkę Petriego (w postaci tzw. przecinków) - inkubacja 37ºC, 16 godz.
PIERWSZY TEST WERYFIKACYJNY		
przygotowanie matrycy dla reakcji PCR	próbówki eppendorf z wodą (50 µL)	- pobranie próbki, czyli kilku komórek z wybranego przecinka (kolonii) - rozprowadzenie próbki w wodzie - dezintegracja komórek, 95ºC, 10 min. - schłodzenie próbki na lodzie, 5 min. - zwirowanie próbki, 3 min., 14000 rpm
reakcja PCR	polimeraza Taq (Fermentas) komercyjne startery przeznaczone do sekwencjonowania insertów danego wektora	- przeprowadzenie standardowej reakcji zawierającej ok. 30 cykli amplifikacji
ocena produktów reakcji PCR	- elektroforeza w 1% żelu agarozowym	
DRUGI TEST WERYFIKACYJNY		
hodowla na pożywce selektywnej	pożywka płynna LB z ampicyliną (0.1 mg/mL)	- pobranie próbki, czyli kilku komórek z wybranego przecinka (kolonii) - zaszczepienie świeżej pożywki LB (3 mL) - inkubacja 37ºC, 16 godz., 250 rpm
izolacja plazmidowych wektorów ekspresyjnych	- wykonana zgodnie z protokołem literaturowym (**Sambrook i Russel, 2001;** Rozdział 1, Protokół nr 1.32) - lub przy użyciu komercyjnego zestawu i protokołu: Plasmid DNA purification using the QIAprep spin miniprep kit and microcentrifuge (Qiagen)	
sekwencjonowanie	- wykonane przez firmę GENOMED z Warszawy	
zabezpieczenie wektorów	- zamrożenie (-20ºC) tych wektorów, które pozytywnie przeszły oba testy weryfikacyjne	



Rysunek 5.3. Przykładowy wynik pierwszego testu weryfikacyjnego dla pMAL-wm9-NodA. Żel agarozowy zawierający produkty reakcji PCR amplifikowane na matrycy z różnych kolonii. Ścieżki na żelu agarozowym: 1, 3, 4, 5, 6 – zawierają fragment długości 842 p.z., czyli insert NodA (637 p.z.) + dwa (102 i 103 p.z.) fragmenty wektora, te kolonie poddawałam kolejnemu testowi weryfikacyjnemu; ścieżki 2 i 7 – to produkty PCR świadczące o nieprawidłowym klonowaniu; ścieżka 8 – marker DNA.

5.2.3. Pilotażowa ekspresja białek nodulacji

Przygotowane przeze mnie plazmidowe wektory ekspresyjne (Tabela 5.2) były kolejno poddawane pilotażowym testom ekspresji, podczas których oceniałam ilość i jakość nadprodukowanych białek nodulacji. Sprawdzałam, czy mają odpowiednią, oczekiwaną masę oraz czy są rozpuszczalne. Wszystkie testy obejmowały te same cztery etapy (transformacja do komórek gospodarza, hodowla komórek, ekspresja białka, analiza wyników), choć w szczegółach znacznie się od siebie różniły. Ekspresję poszczególnych białek nodulacji prowadziłam nie tylko z różnych wektorów, ale także w różnych komórkach ekspresyjnych, różnych pożywkach i temperaturach (Tabela 5.5. i Rysunek 5.4). Sprawdzałam również, czy czas prowadzenia ekspresji wpływa na obecność rekombinowanych białek we frakcjach rozpuszczalnych i nierozpuszczalnych.

5.2.3.1. Transformacja komórek gospodarza

Wprowadzając dany wektor do komórki gospodarza *E. coli* posługiwałam się szokiem temperaturowym. Próbkę zawierającą wektor i komórki inkubowałam kolejno na lodzie (30 min.), w bloku grzejnym o temperaturze 42ºC (20 - 60 sek.) i ponownie na lodzie (5 min.). Do transformacji wykorzystywałam różne komórki ekspresyjne: BL21 Star, NEB express, BL magic oraz podnoszące wydajność ekspresji komórki OverExpress C41 i C43 (**Miroux i Walker, 1996**).

5.2.3.2. Hodowla komórek

Po transformacji prowadziłam hodowle komórek gospodarza *E. coli*. Trzy następujące po sobie hodowle bakterii, każda zaszczepiana komórkami poprzedniej, prowadzone były w 37ºC, kolejno:
- przez 1 godzinę, z wytrząsaniem (200 rpm), w pożywce SOC (200 μL),
- przez noc (ok. 16 godz.), na pożywce stałej LB z antybiotykiem,
- przez 2 - 4 godz., z wytrząsaniem (250 rpm), w pożywce z antybiotykiem (10 mL).

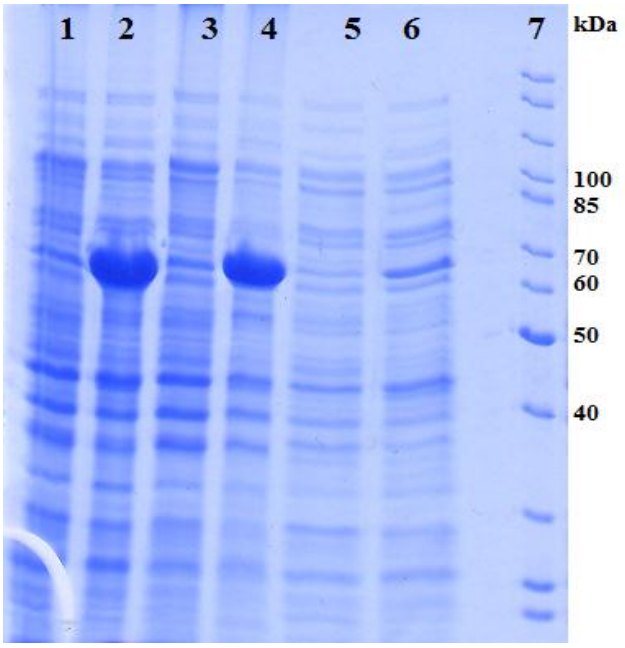
Ostatnia hodowla płynna była prowadzona do osiągnięcia wartości OD_{600} w granicach 0.5 - 0.8 i odbywała się w różnego rodzaju pożywkach (LB, TB, minimalnej), których skład mógł wpłynąć na rozpuszczalność eksprymowanych białek (**Moore *et al.*, 1993**).

5.2.3.3. Ekspresja białek

Kulturę bakterii dzieliłam na dwie równe części (po 5 mL), z których jedna służyła do testów ekspresji (indukowana), a druga stanowiła próbę kontrolną (nieindukowana). Ekspresję indukowałam za pomocą IPTG, przy czym stosując różne stężenia końcowe (w zakresie od 0.1 mM do 0.5 mM) kontrolowałam poziom produkcji rekombinowanych białek. Aby ułatwić fałdowanie się rekombinowanych białek, indukowałam również ekspresję białek opiekuńczych, konkretnie białek szoku cieplnego (*ang. heat shock proteins*), naturalnie występujących w bakteriach *E. coli* (**Arsene *et al.*, 2000**). W tym celu na 30 - 60 minut podwyższałam temperaturę hodowli do 42ºC, następnie obniżałam ją do temperatury ekspresji i dopiero potem indukowałam produkcję białka nodulacji. Ekspresje prowadziłam w trzech różnych temperaturach: 18ºC, 27ºC, 37ºC, ponieważ temperatura ekspresji potrafi wpływać na rozpuszczalność produkowanych białek (**Schein i Noteborn, 1988**).

5.2.3.4. Analiza wyników ekspresji

W trakcie każdego pilotażowego testu ekspresji, z obu kultur, indukowanej i nieindukowanej, pobierałam kilka do kilkunastu próbek. Próbki te odpowiednio preparowałam (poddawałam lizie i rozdzielałam rozpuszczalne i nierozpuszczalne składniki komórki poprzez wirowanie), a ich zawartość analizowałam za pomocą elektroforezy SDS-PAGE.



Rysunek 5.4. Pilotażowa ekspresja białka NodA (65.9 kDa) z wektora pMAL-wm9-NodA wykonana w komórkach NEB, w pożywce LB, w temperaturze 18ºC, indukowana IPTG (0.3 mM) po ekspresji białek opiekuńczych (42ºC). Ścieżki na żelu SDS-PAGE:
1- kultura nieindukowana, wszystkie białka,
2- kultura indukowana, wszystkie białka,
3-kultura nieindukowana, białka nierozpuszczalne
4- kultura indukowana, białka nierozpuszczalne,
5- kultura nieindukowana, białka rozpuszczalne,
6- kultura indukowana, białka rozpuszczalne,
7 – Marker masy białek.

Tabela. 5.5. Pilotażowe ekspresje białek nodulacji. Wyszczególniam, jakie warunki ekspresji były testowane dla poszczególnych wektorów oraz podsumowuję wyniki.

wektor ekspresyjny	testowane warunki ekspresji	wyniki ekspresji pilotażowych
pET-wm9-NodA	komórki: BL21Star pożywki: LB, TB, minimalna IPTG: 0.1 mM, 0.3 mM, 0.5 mM temperatura: 18ºC, 27ºC, 37ºC	białko NodA ulegało ekspresji głównie w formie nierozpuszczalnej nieznaczną pulę białek (5 - 20%) w formie rozpuszczalnej uzyskiwałam : - prowadząc ekspresję w 18ºC po indukcji białek opiekuńczych - produkując białko NodA w fuzji z MBP pozostałe zmienne (komórki, pożywki, stężenie IPTG) nie wpływały na ekspresję lub miały drugorzędne znaczenie (nieznacznie zmieniając proporcję białek rozpuszczalnych i nierozpuszczalnych)
pET-mel-NodA	komórki: BL21Star, C41, C43 pożywki: LB, TB, minimalna IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC, 27ºC, 37ºC indukcja białek opiekuńczych (42ºC)	
pMAL-wm9-NodA	komórki: BL21Star, C41, C43, NEB pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC, 27ºC, 37ºC indukcja białek opiekuńczych (42ºC)	
pMB-wm9-NodA	komórki: BL21Star, BL21magic pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC indukcja białek opiekuńczych (42ºC)	
pET-wm9-NodU	komórki: BL21Star pożywki: LB, TB, minimalna IPTG: 0.1 mM, 0.3 mM, 0.5 mM temperatura: 18ºC, 27ºC, 37ºC	białko NodU ulegało ekspresji tylko i wyłącznie w formie nierozpuszczalnej niezależnie od fizyko-chemicznych warunków ekspresji, czy charakterystyki elementów systemu ekspresyjnego (komórek i wektorów)
pET-ngr-NodU	komórki: BL21Star pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC, 37ºC indukcja białek opiekuńczych (42ºC)	
pMAL-wm9-NodU	komórki: BL21Star, C41, C43, NEB pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC, 37ºC indukcja białek opiekuńczych (42ºC)	

Tabela 5.5. c.d.

pET-wm9-NolO	komórki: BL21Star, C41, C43 pożywki: LB, TB, minimalna IPTG: 0.1 mM, 0.3 mM, 0.5 mM temperatura: 18ºC, 27ºC, 37ºC indukcja białek opiekuńczych (42ºC)	białko NolO ulegało ekspresji tylko i wyłącznie w formie nierozpuszczalnej niezależnie od fizyko-chemicznych warunków ekspresji, czy charakterystyki elementów systemu ekspresyjnego (komórek i wektorów)
pET-ngr-NolO	komórki: BL21Star pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC, 37ºC indukcja białek opiekuńczych (42ºC)	
pMAL-wm9-NolO	komórki: BL21Star, C41, C43, NEB pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC, 37ºC indukcja białek opiekuńczych (42ºC)	
pET-wm9-NodB	komórki: BL21Star, C41, C43 pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC indukcja białek opiekuńczych (42ºC)	białko NodB ulegało ekspresji głównie w formie nierozpuszczalnej nieznaczną pulę białek w formie rozpuszczalnej (ok. 10%) uzyskiwałam jedynie gdy białko NodB znajdowało się w fuzji z MBP, w wersji z krótkim łącznikiem, przy czym konieczna była równoczesna indukcja białek opiekuńczych i ekspresja w 18ºC pozostałe zmienne (komórki, pożywki, stężenie IPTG) nie wpływały na ekspresję
pMAL-wm9-NodB	komórki: BL21Star pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC indukcja białek opiekuńczych (42ºC)	
pMB-wm9-NodB	komórki: BL21magic pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC indukcja białek opiekuńczych (42ºC)	
pET-jap-NoeI	komórki: BL21Star, C41, C43 pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC indukcja białek opiekuńczych (42ºC)	białko NoeI ulegało ekspresji głównie w formie nierozpuszczalnej nieznaczną pulę białek w formie rozpuszczalnej (ok. 10%) uzyskiwałam jedynie gdy białko NoeI znajdowało się w fuzji z MBP, w wersji z krótkim łącznikiem, przy czym konieczna była równoczesna indukcja białek opiekuńczych i ekspresja w 18ºC pozostałe zmienne (komórki, pożywki, stężenie IPTG) nie wpływały na ekspresję
pMAL-jap-NoeI	komórki: BL21Star pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC indukcja białek opiekuńczych (42ºC)	
pMB-jap-NoeI	komórki: BL21magic pożywki: TB IPTG: 0.1 mM, 0.3 mM temperatura: 18ºC indukcja białek opiekuńczych (42ºC)	

5.2.4. Produkcja, oczyszczanie i krystalizacja wybranych białek nodulacji

Przed przystąpieniem do ekspresji na większą skalę (nadprodukcji) każde kolejne białko nodulacji wstępnie analizowałam bioinformatycznie za pomocą narzędzia ProtParam z serwera Expasy (**Gasteiger *et al.*, 2005**). Dzięki temu uzyskiwałam orientacyjne informacje przydatne podczas procesu oczyszczania danego białka.

5.2.4.1. NodA z *Bradyrhizobium* sp. WM9 ze znacznikiem histydynowym

Białko NodA z *Bradyrhizobium* sp. WM9 ulegające ekspresji z wektora pET-wm9-NodA na N-końcu posiadało znacznik histydynowy i łączną masę 27.4 kDa. Istniała możliwość odcięcia znacznika za pomocą proteazy TEV, wtedy masa białka zmniejszyłaby się do 23.4 kDa. Miejsce cięcia było tak usytuowane, że na aminowym końcu białka pozostawiało sześć dodatkowych aminokwasów, nieobecnych w natywnej sekwencji. Teoretyczna wartość pI białka wynosi 9.5 sugerując, że to białko silnie zasadowe. Sekwencja NodA zawiera jedną resztę cysteiny, dlatego ewentualne mostki dwusiarczkowe mogłyby się tworzyć tylko w kontekście międzycząsteczkowym.

Produkcja białka

Ta wersja białka NodA (oznaczona skrótem NodA-wm9) ulegała ekspresji tylko i wyłącznie w formie nierozpuszczalnej, niezależnie od zastosowanych warunków. Postanowiłam zatem podjąć próbę uzyskania preparatu z ciałek inkluzyjnych. Białko do tego celu było produkowane zgodnie z protokołem zamieszczonym w Tabeli 5.6.

Tabela 5.6. Protokół produkcji rekombinowanego białka NodA-wm9.

etap	stosowane materiały	procedura
transformacja	wektor pET-wm9-NodA komórki BL21 Star (INVITROGEN) pożywka płynna SOC	- delikatne zmieszanie rozmrożonych komórek z wektorem (ok. 5 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5 min. - zmieszanie komórek z SOC (250 µL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla na pożywce stałej	pożywka stała LB z ampicyliną (0.1 mg/mL)	- posiew na płytkę z pożywką stałą LB - inkubacja 37ºC, 16 godz.
hodowla na pożywce płynnej	pożywka płynna TB z ampicyliną (0.1 mg/mL)	- przeniesienie komórek z płytki do pożywki płynnej TB (250 mL) - inkubacja 37ºC, 2 godz., 250 rpm - rozcieńczenie hodowli świeżą pożywką TB (do 2 L) - inkubacja 37ºC, 250 rpm, do OD_{600}=0.8 (ok 3 godz.) - obniżenie temperatury do 18ºC
ekspresja	IPTG (do stężenia 0.3 mM)	- dodanie IPTG - inkubacja 18ºC, 16 godz., 250 rpm
zabezpieczenie komórek		- wirowanie 5000 rmp, 10ºC, 30 min. - zamrożenie komórek w -80ºC

Obróbka ciałek inkluzyjnych

Z komórek zawierających nadprodukowane nierozpuszczalne białko NodA-wm9 izolowałam ciałka inkluzyjne, które rozpuszczałam za pomocą mocznika (Tabela 5.7). Następnie przechodziłam do obróbki uzyskanej mieszaniny rozfałdowanych białek, której celem było doprowadzenie do uzyskania rozpuszczalnej formy natywnej białka NodA-wm9. Ponowne fałdowanie białka przeprowadzałam usuwając czynnik denaturujący, przy czym używałam nieoczyszczonych, wstępnie oczyszczonych lub związanych do złoża niklowego białek NodA-wm9.

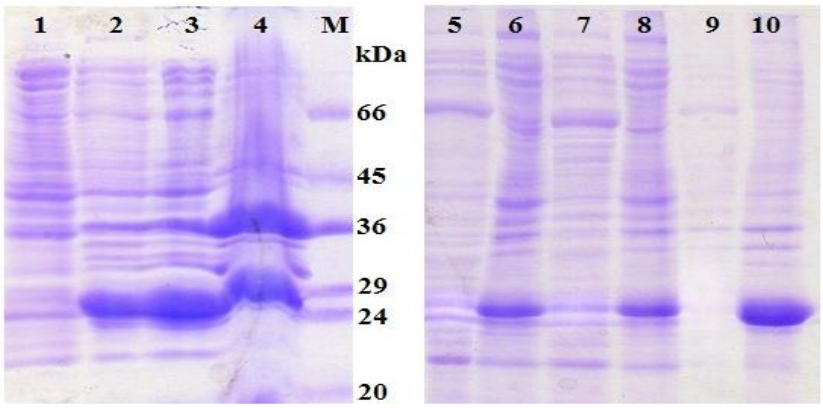
Tabela 5.7. Protokół obróbki ciałek inkluzyjnych zawierających NodA-wm9.

etap	stosowane materiały	procedura
liza komórek	roztwór 1: 50 mM Tris pH 8.0 10 mM EDTA 1 mM PMSF 5 mg/mL lizozym	- rozpuszczenie osadu komórkowego w roztworze 1 (ok. 50 mL / 4 g osadu) - 0.5 godz. delikatnego mieszania w 20ºC - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - wirowanie: 14000 rpm, 10ºC, 30 min. - odrzucenie supernatantu
rozpuszczanie ciałek inkluzyjnych	roztwór 2: 50 mM Tris pH 8.0 5% glicerol 5 mM β-merkaptoetanol 7 M mocznik	- rozpuszczenie osadu w roztworze 2 (ok. 10 mL) - 1 godz. delikatnego mieszania w 37ºC - wirowanie: 14000 rpm, 10ºC, 3 godz. - odrzucenie osadu

Ponowne fałdowanie nieoczyszczonego białka

Mieszaninę rozfałdowanych mocznikiem białek poddawałam powolnej kilkuetapowej dializie do roztworów o skokowo zmniejszającym się stężeniu czynnika denaturującego. Kolejno stosowane roztwory zawierały 5 M, 3 M, 2 M, 1 M, 0.5 M i 0 M mocznik. Pozostałe składniki roztworów, dla osobnych eksperymentów, dobierałam tak aby jak najbardziej zróżnicować warunki renaturacji białek. Przetestowałam różne rodzaje i pH buforów (MES pH 6, Tris pH 7 i pH 8, fosforanowy pH 7.5, CAPS pH 10) oraz rodzaje i stężenia soli (NaCl i KCl: 500 mM, 200 mM, 0 mM). Stałymi składnikami roztworów były 5% glicerol i 5 mM β-merkaptoetanol. Dializę wykonywałam w 4ºC. Po zakończonej dializie preparat, w którym prawie zawsze widoczne były nitkowate sedymentujące cząstki, wirowałam (20000 rpm, 4ºC, 0.5 godz.) celem usunięcia nierozpuszczalnych białek. Zawartość supernatantu i osadu weryfikowałam na żelach SDS-PAGE (Rysunek 5.5). Niestety fałdowanie nieoczyszczonego białka NodA-wm9 nie powiodło się, gdyż za każdym razem białko strącało się podczas dializy.

Rysunek 5.5. Izolacja, rozpuszczenie i ponowne fałdowanie nieoczyszczonych białek NodA-wm9 ulegających ekspresji w formie nierozpuszczalnej (w ciałkach inkluzyjnych). Ścieżki na żelu SDS-PAGE:
1- białka rozpuszczalne uwolnione z komórki; 2- białka nierozpuszczalne uwolnione z komórki;
3-4- efekty rozpuszczania ciałek inkluzyjnych mocznikiem: 3- rozfałdowane białka zawarte w supernatancie, 4- białka, których nie udało się rozpuścić zawarte w osadzie; M- marker masy;
5-10- fałdowanie nieoczyszczanego białka NodA-wm9 za pomocą dializy do roztworu o końcowym składzie: 20 mM bufor fosforanowy pH 7.5, 5% glicerol, 5 mM β-merkaptoetanol, 0 - 500 mM NaCl:
5-6- roztwór dializacyjny zawierający 500 mM NaCl, 5- białka rozpuszczalne, 6- białka nierozpuszczalne
7-8- roztwór dializacyjny zawierający 200 mM NaCl, 7- białka rozpuszczalne, 8- białka nierozpuszczalne
9-10- roztwór dializacyjny nie zawierający NaCl, 9- białka rozpuszczalne, 10- białka nierozpuszczalne.



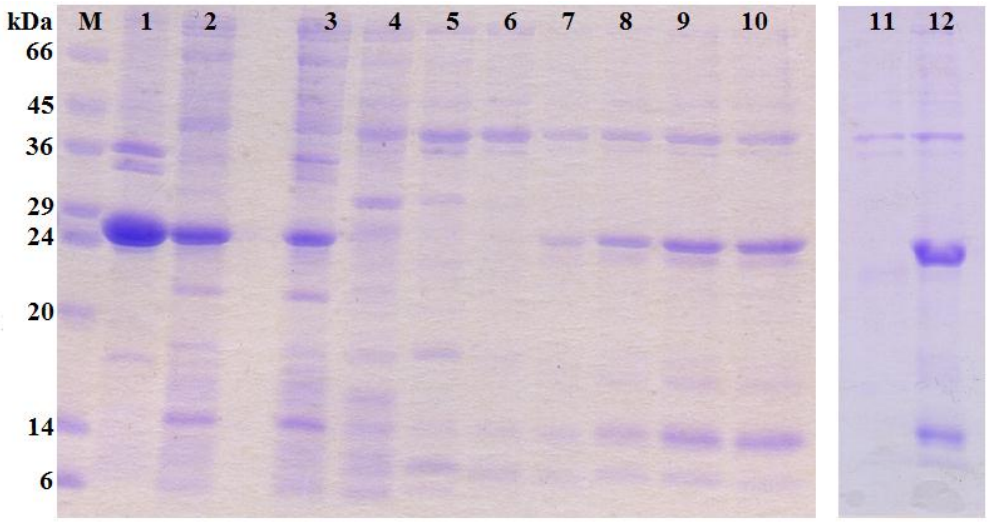
Ponowne fałdowanie wstępnie oczyszczonego białka

Z mieszaniny rozfałdowanych mocznikiem białek wyodrębniałam NodA-wm9 przy użyciu chromatografii powinowactwa, którą przeprowadziłam w warunkach denaturujących (Tabela 5.8 i Rysunek 5.6).

Tabela 5.8. Chromatografia powinowactwa wykonana w warunkach denaturujących, w temperaturze 4ºC metodą grawitacyjną przy użyciu złoża niklowego His-Bind Resins (Novagen) załadowanego $NiSO_4$.

etap	skład stosowanych roztworów	procedura
wiązanie do złoża	roztwór 1: 50 mM bufor fosforanowy pH 8.0 500 mM NaCl 5% glicerol 5 mM β-merkaptoetanol 7 M mocznik	- wypełnienie kolumny złożem (ok. 10 mL) - przemywanie złoża roztworem 1 - dwukrotne przepuszczenie przez złoże mieszaniny rozfałdowanych białek - odrzucenie eluatu zawierającego frakcję białek niezwiązanych do złoża
przemywanie	roztwór 2: 50 mM bufor fosforanowy pH 6.0 500 mM NaCl 5% glicerol 5 mM β-merkaptoetanol 7 M mocznik	- przemywanie złoża roztworem 2 (ok. 50 mL)
wymywanie	roztwór 3: 50 mM bufor fosforanowy pH 4.0 500 mM NaCl 5% glicerol 5 mM β-merkaptoetanol 7 M mocznik	- wymywanie związanych do złoża białek roztworem 3 (ok. 10 mL x 5)

Rysunek 5.6. Chromatografia powinowactwa wykonana w warunkach denaturujących oraz wynik fałdowania wstępnie oczyszczonych białek. Ścieżki na żelu SDS-PAGE: M- marker masy; 1- białka nierozpuszczalne uwolnione z komórki (ciałka inkluzyjne); 2- rozfałdowane białka pochodzące z ciałek inkluzyjnych przeznaczone do oczyszczania; 3-10- chromatografia powinowactwa: 3- frakcja białek nie związanych do złoża; 4-5- frakcje przemywania; 6-10- kolejne frakcje zawierające białka wymywane ze złoża; 11-12- fałdowanie oczyszczonego białka NodA-wm9 za pomocą dializy, 11- białka rozpuszczalne (supernatant), 12- białka nierozpuszczalne (osad).



Po chromatografii powinowactwa, frakcje zawierające wstępnie oczyszczone, nadal rozfałdowane białko NodA-wm9 łączyłam i następnie poddawałam powolnej kilkuetapowej dializie przeprowadzonej dokładnie tak jak dializy białka nie-oczyszczonego. Tak jak wcześniej, podczas osobnych eksperymentów testowałam różne warunki renaturacji. Po zakończonej dializie preparat wirowałam (20000 rpm, 4ºC, 0.5 godz.), a uzyskane próbki supernatantu (białek rozpuszczanych) i osadu (białek nierozpuszczalnych) weryfikowałam na żelach SDS-PAGE (Rysunek 5.6). Niestety próba fałdowania oczyszczonego białka NodA-wm9 również nie powiodła się.

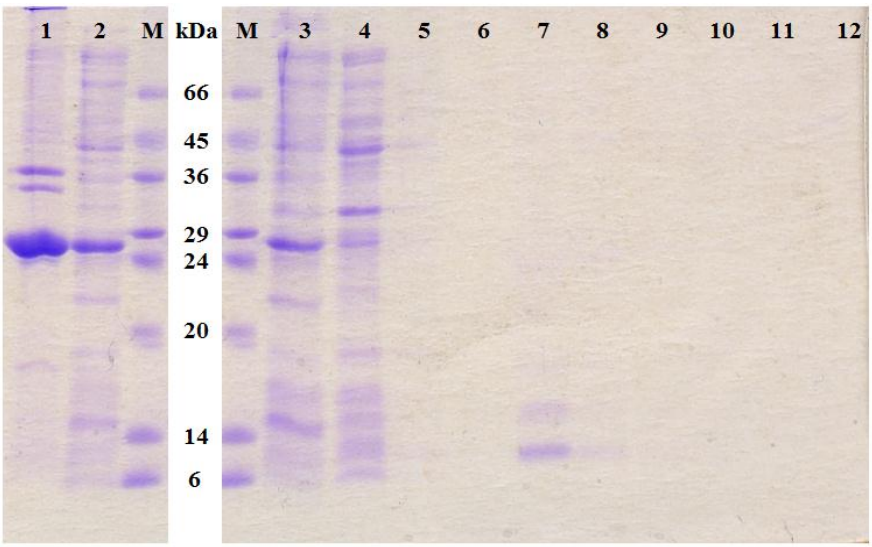
Ponowne fałdowanie białka związanego do złoża

Podjęłam także próbę ponownego fałdowania białka NodA-wm9 podczas chromatografii powinowactwa w tzw. warunkach hybrydowych (Tabela 5.9 i Rysunek 5.7). Z mieszaniny rozpuszczonych mocznikiem białek wyodrębniałam NodA-wm9 wiążąc je do złoża niklowego. Podczas wiązania wszystkie białka pozostawały rozfałdowane. Następnie stopniowo zmieniłam warunki chromatografii z denaturujących na natywne i rozpoczęłam wymywanie związanych do złoża (ponownie sfałdowanych) białek. Niestety białko NodA-wm9 najprawdopodobniej strąciło się podczas renaturacji i utknęło na kolumnie.

Tabela 5.9. Chromatografia powinowactwa wykonana w warunkach hybrydowych, prowadzona w temperaturze 4ºC metodą grawitacyjną przy użyciu złoża niklowego His-Bind Resins (Novagen) załadowanego $NiSO_4$.

etap	skład stosowanych roztworów	procedura
wiązanie do złoża	roztwór 1: 50 mM bufor fosforanowy pH 8.0 500 mM NaCl 5% glicerol 5 mM β-merkaptoetanol 7 M mocznik	- wypełnienie kolumny złożem (ok. 10 mL) - przemywanie złoża roztworem 1 - dwukrotne przepuszczenie przez złoże mieszaniny rozfałdowanych białek - odrzucenie eluatu zawierającego frakcję białek niezwiązanych do złoża
przemywanie	roztwór 2: 50 mM bufor fosforanowy pH 6.0 500 mM NaCl 5% glicerol 5 mM β-merkaptoetanol 7 M mocznik	- przemywanie złoża roztworem 2 (ok. 50 mL)
fałdowanie białek	roztwór 3: 50 mM bufor fosforanowy pH 8.0 500 mM NaCl 5% glicerol 5 mM β-merkaptoetanol	- wielokrotne przemywanie złoża kolejnymi mieszaninami roztworów: 1 (5 mL) i 3 (0 mL) 1 (4 mL) i 3 (1 mL) 1 (3 mL) i 3 (2 mL) 1 (2 mL) i 3 (3 mL) 1 (1 mL) i 3 (4 mL) 1 (0 mL) i 3 (5 mL)
przemywanie	roztwór 4: roztwór 3 + 20 mM imidazol	- przemywanie złoża roztworem 4 (ok. 50 mL)
wymywanie	roztwór 5: roztwór 3 + 300 mM imidazol roztwór 6: roztwór 3 + 500 mM imidazol roztwór 7: roztwór 3 + 1 M imidazol	- wymywanie związanych do złoża białek kolejnymi roztworami: 5 (ok. 10 mL x 2) 6 (ok. 10 mL x 2) 7 (ok. 10 mL x 2)

Rysunek 5.7. Chromatografia powinowactwa wykonana w warunkach hybrydowych. Ścieżki na żelu SDS-PAGE: M- marker masy; 1- białka nierozpuszczalne (ciałka inkluzyjne); 2- rozfałdowane białka pochodzące z ciałek inkluzyjnych przeznaczone do oczyszczania; 3-12- chromatografia powinowactwa: 3- frakcja białek nie związanych do złoża; 4-5- frakcja przemywania przed fałdowaniem; 6- frakcja przemywania po fałdowaniu; 7-12- kolejne frakcje zawierające białka wymywane ze złoża.



5.2.4.2. NodA z *Sinorhizobium meliloti* ze znacznikiem histydynowym

Białko NodA z *Sinorhizobium meliloti* 1021 ulegające ekspresji z wektora pET-mel-NodA na N-końcu posiadało znacznik histydynowy, który można było odciąć za pomocą proteazy TEV, pozostawiając na białku sześć dodatkowych aminokwasów, nieobecnych w natywnej sekwencji. Oczekiwana masa białka po ekspresji wynosiła 25.8 kDa, a po odcięciu znacznika 21.8 kDa. Teoretyczne pI tej wersji białka NodA wynosi 8.7. Posiada ono 2 reszty cysteiny w sekwencji. Sekwencje NodA z *Sinorhizobium meliloti* i *Bradyrhizobium* sp. WM9 znacznie się od siebie różnią (identyczność 63%, podobieństwo 78%).

Produkcja białka

Pilotażowe testy ekspresji tej wersji NodA (oznaczonej skrótem NodA-mel) wykazały, że białka opiekuńcze obecne podczas produkcji pomagają uzyskać rozpuszczalny produkt (Tabela 5.10). Oszacowałam, że stanowi on zaledwie 5 - 10% całej puli nadprodukowanego białka NodA-mel, jednak pozwala na podjęcie próby oczyszczania z frakcji białek rozpuszczalnych.

Tabela 5.10. Protokół produkcji rekombinowanego białka NodA-mel.

etap	stosowane materiały	procedura
transformacja	wektor pET-mel-NodA komórki BL21 Star (INVITROGEN) pożywka płynna SOC	- delikatne zmieszanie rozmrożonych komórek z wektorem (ok. 5 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5 min. - zmieszanie komórek z SOC (250 μL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla na pożywce stałej	pożywka stała LB z ampicyliną (0.1 mg/mL)	- posiew na płytkę z pożywką stałą LB - inkubacja 37ºC, 16 godz.
hodowla na pożywce płynnej	pożywka płynna TB z ampicyliną (0.1 mg/mL)	- przeniesienie komórek z płytki do pożywki płynnej TB (250 mL) - inkubacja 37ºC, 2 godz., 250 rpm - rozcieńczenie hodowli świeżą pożywką TB (do 2 L) - inkubacja 37ºC, 250 rpm, do OD_{600}=0.8 (ok. 3godz.)
indukcja białek opiekuńczych		- inkubacja 42ºC, 250 rpm, 1 godz. - obniżenie temperatury do 18ºC
ekspresja	IPTG (do stężenia 0.1 mM)	- dodanie IPTG - inkubacja 18ºC, 16 godz., 250 rpm
zabezpieczenie komórek		- wirowanie 5000 rmp, 10ºC, 30 min. - zamrożenie komórek w -80ºC

Oczyszczanie białka

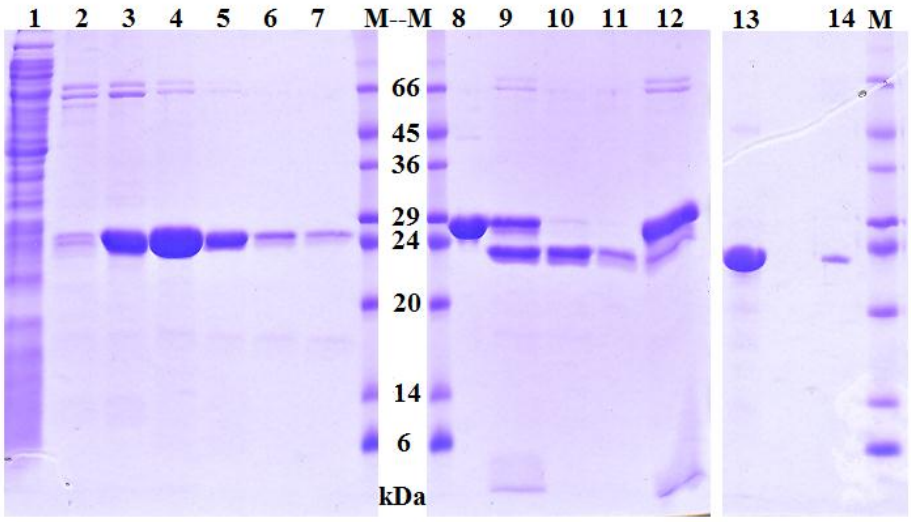
Wielokrotnie oczyszczając białko NodA-mel, za każdym razem postępowałam nieco inaczej, modyfikując procedurę przez dodawanie lub pomijanie etapów oraz stosując różne odczynniki. Postępowałam tak z dwóch powodów. Po pierwsze chciałam poprawić wydajność izolacji, która była bardzo niska, ze względu na niską ekspresję białek rozpuszczalnych. Po drugie starałam się zwiększyć stabilność białka NodA-mel, które bardzo łatwo wytrącało się z roztworu, często wręcz uniemożliwiając dokończenie procedury oczyszczania.

Tabela 5.11. Przykładowa procedura oczyszczania białka NodA-mel.

etap	stosowane materiały	procedura
liza	roztwór 1: 50 mM Tris pH 7.5 500 mM NaCl 5 mM β-merkaptoetanol 5% glicerol 1.6 mM CHAPS 20 mM imidazol PMSF (do stężenia 1 mM) lizozym (do stężenia 1 mg/mL)	- rozpuszczenie osadu komórkowego w roztworze 1 (ok. 50 mL / 4 g osadu) - dodanie PMSF i lizozymu - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - 1 godz. delikatnego mieszania w 4ºC - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - wirowanie: 14000 rpm, 4ºC, 30 min. - odrzucenie osadu
chromatografia powinowactwa na kolumnie niklowej metodą grawitacyjną	złoże His-Bind Resins (Novagen) załadowane $NiSO_4$ i przemyte roztworem 1 roztwór 2: 50 mM Tris pH 7.5 500 mM NaCl 5 mM β-merkaptoetanol 5% glicerol 1.6 mM CHAPS 0.1 - 1 M imidazol cały etap prowadzony w temperaturze 4ºC	- zmieszanie supernatantu ze złożem wypełniającym kolumnę (ok. 10 mL złoża / 50 mL supernatantu) - inkubacja 30 min. - odrzucenie eluatu zawierającego frakcję białek niezwiązanych do złoża - przemywanie roztworem 1 (ok. 20 mL) - wymywanie białka roztworem 2 (ok. 20 mL) zawierającym stopniowo zwiększające się stężenie imidazolu - połączenie frakcji zawierających NodA-mel
usunięcie znacznika histydynowego	TEV^{His} (0.15 mg / 1 mg białka NodA-mel)	- zmieszanie eluatu z proteazą TEV^{His} - dializa do roztworu 1 (16 godz., 4ºC, delikatne mieszanie)
chromatografia powinowactwa na kolumnie niklowej metodą grawitacyjną	złoże His-Bind Resins (Novagen) załadowane $NiSO_4$ i przemyte roztworem 1 cały etap prowadzony w temperaturze 4ºC	- zmieszanie przedializowanych białek ze złożem kolumny (ok. 10 mL złoża / 20 mL roztworu) - inkubacja 30 min. - zebranie eluatu zawierającego białko NodA-mel - zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (Millipore) - odrzucenie białek, które pozostały na złożu

Najczęściej oczyszczanie prowadziłam stosując protokół zawierający cztery etapy: lizy, chromatografii powinowactwa, usuwania znacznika histydynowego i ponownej chromatografii powinowactwa (Tabela 5.11 i Rysunek 5.8). Protokół ten pozwalał szybko, w przeciągu dwóch dni, uzyskać preparat gotowy do krystalizacji, co miało szczególne znaczenie w przypadku łatwo strącającego się białka NodA-mel. Próbowałam jeszcze bardziej przyspieszyć procedurę, poprzez skrócenie czasu trwania szesnastogodzinnego etapu usuwania znacznika histydynowego. Zamiast dializy, niezbędnej do usunięcia imidazolu przed ponowną chromatografią powinowactwa, zastosowałam kolumnę odsalająca NAP-10 (GE-Healthcare), natomiast proces cięcia proteazą TEV^{His} przyspieszyłam podwyższając temperaturę reakcji do 30ºC. Jednak drastyczna zmiana warunków podczas odsalania oraz wysoka temperatura cięcia spowodowały jeszcze szybsze wytrącanie się białka NodA-mel. Procedurę modyfikowałam również poprzez dodanie etapów strącania siarczanem amonu, chromatografii jonowymiennej, czy chromatografii żelowej, jednak metody te nie dawały dobrych rezultatów, a jedynie niepotrzebnie wydłużały procedurę. Ważną zaletą czteroetapowego protokołu było to, że zatężanie białka, podczas którego następowały największe straty, wykonywałam tylko raz.

Rysunek 5.8. Wynik przykładowej procedury oczyszczania białka NodA-mel. Ścieżki na żelu SDS-PAGE: 1-7- etap chromatografii powinowactwa: M- marker masy; 1- frakcja białek nie związanych do złoża, 2-7-wymywanie białka ze złoża niklowego zwiększającym się stężeniem imidazolu; 8-9- etap odcinania znacznika histydynowego: 8- proteaza TEV^{His}; 9- białka zawarte w preparacie po etapie odcinania znacznika; 10-12- etap drugiej chromatografii powinowactwa: 10- eluat białek NodA-wm9 pozbawionych znacznika które nie związały się do złoża, 11- przemywanie złoża, 12- białka które pozostały związane do złoża, 13- zatężony preparat NodA-mel, 14- rozcieńczony preparat NodA-mel. Za pomocą spektrometrii mass zidentyfikowano białko zawarte w prążku ścieżki 14 jako NodA z *Sinorhizobium meliloti*.



W trakcie izolacji i oczyszczania białka NodA-mel testowałam różne odczynniki, mając nadzieję na znalezienie takiego, który zapobiegnie strącaniu się preparatu. Stosowałam różne bufory (Tris pH 9.0 i pH 7.5 oraz fosforanowy pH 6.0), sole i ich stężenia (NaCl i KCl: 0 - 500 mM) oraz wiele rozmaitych dodatków do roztworów (5 mM β-merkaptoetanol, 5 - 10% glicerol, 0.005 - 0.5% TritonX100, 0.01 - 0.1% Tween20, 1.6 - 15 mM CHAPS, 23 - 40 mM oktyloglukoza, 200 mM NDSB201, 200 mM TMAO, 10 mM $MgSO_4$) (**Bondos i Bicknell, 2003; Prive, 2007**).

Krystalizacja białka

Najlepszy wynik oczyszczania jaki uzyskałam to ok. 2 mg białka z 1 L hodowli, przy czym maksymalne stężenie preparatu wynosiło ok. 4 mg/mL, niezależnie od składu roztworu, w którym preparat otrzymałam. Białko NodA-mel poddawałam krystalizacji korzystając z komercyjnych zestawów do krystalizacji firm Hampton i MDL. Nastawiałam krople wiszące (ręcznie) w trzech w proporcjach preparatu do rezerwuaru (1:3, 2:2, 3:1) i pozostawiałam w temperaturze 19ºC lub 4ºC. Niestety, za każdym razem, już następnego dnia po nastawianiu krystalizacji wszystkie krople zawierały strącone białko.

5.2.4.3. NodA z *Bradyrhizobium* sp. WM9 z białkiem fuzyjnym MBP (wersja 1.)

Białko NodA z *Bradyrhizobium* sp. WM9 ulegające ekspresji z wektora pMAL-wm9-NodA miało znacznie większą masę od swojej natywnej wersji z uwagi na obecność białka fuzyjnego MBP (65.9 kDa). Istniała możliwość oddzielenia białek NodA (23.4 kDa) i MBP (42.5 kDa) poprzez cięcie proteazą - czynnikiem Xa. Miejsce cięcia było tak zaprojektowane, żeby nie pozostawiać żadnych dodatkowych reszt na białku nodulacji, a dwudziestoaminokwasowy łącznik pozostawał przy MBP.

Produkcja białka

Ta wersja białka NodA (oznaczona skrótem NodA-20-MBP) również ulegała ekspresji zarówno w formie rozpuszczalnej, jak i nierozpuszczalnej. Pilotażowe testy ekspresji pozwoliły określić warunki, w których pula nadprodukowanych rozpuszczalnych białek była największa (do 20%). Rozpuszczalne białko przeznaczone do

oczyszczania było produkowane zgodnie z protokołem zamieszczonym w Tabeli 5.12.

Tabela 5.12. Protokół produkcji rekombinowanego białka NodA-20-MBP.

etap	stosowane materiały	procedura
transformacja	wektor pMAL-wm9-NodA komórki BL21 Star (INVITROGEN) pożywka płynna SOC	- delikatne zmieszanie rozmrożonych komórek z wektorem (ok. 5 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5 min. - zmieszanie komórek z SOC (250 µL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla na pożywce stałej	pożywka stała LB z ampicyliną (0.1 mg/mL)	- posiew na płytkę z pożywką stałą LB - inkubacja 37ºC, 16 godz.
hodowla na pożywce płynnej	pożywka płynna TB z glukozą (0.2% w/v) z ampicyliną (0.1 mg/mL)	- przeniesienie komórek z płytki do pożywki płynnej TB (250 mL) - inkubacja 37ºC, 2 godz., 250 rpm - rozcieńczenie hodowli świeżą pożywką TB (do 2 L) - inkubacja 37ºC, 250 rpm, do OD_{600}=0.8 (ok 3 godz.)
indukcja białek opiekuńczych		- inkubacja 42ºC, 250 rpm, 0.5 godz. - obniżenie temperatury do 18ºC
ekspresja	IPTG (do stężenia 0.1 mM)	- dodanie IPTG - inkubacja 18ºC, 16 godz., 250 rpm
zabezpieczenie komórek		- wirowanie 5000 rmp, 10ºC, 30 min. - zamrożenie komórek w -80ºC

Oczyszczanie białka

Najlepszy protokół oczyszczania rozpuszczalnego białka nadprodukowanego z wektora pMAL-wm9-NodA jaki udało mi się opracować, składa się z czterech etapów: lizy, wysalania i odsalania, chromatografii powinowactwa na złożu wiążącym MBP oraz chromatografii żelowej (Tabela 5.13). Optymalny roztwór w którym uzyskiwałam białko zawiera Tris (20 mM, pH 7.5), NaCl (500 mM) oraz glicerol (5 - 10%). Preparat stanowią dwa połączone dwudziestoaminokwasowym łącznikiem białka NodA i MBP (białko NodA-20-MBP) w ilości ok. 8 mg z 1 L hodowli.

W trakcie oczyszczania następują znaczne straty nadprodukowanego białka, których nie udało mi się uniknąć. W szczególności występują one w trakcie zatężania oraz na ostatnim etapie prac, chromatografii żelowej. Na chromatogramie FPLC (Rysunek 5.9) widać, jak skutecznie etap ten umożliwia oddzielenie czystego preparatu NodA-20-MBP (drugi pik) od prawdopodobnie wadliwie pofałdowanych lub zdenaturowanych i zagregowanych form białka NodA, które pozostawały w roztworze

zapewne tylko dzięki związaniu do partnera MBP (pierwszy pik pojawiający się zaraz po wypływie objętości zerowej z kolumny, która wynosi 40 mL).

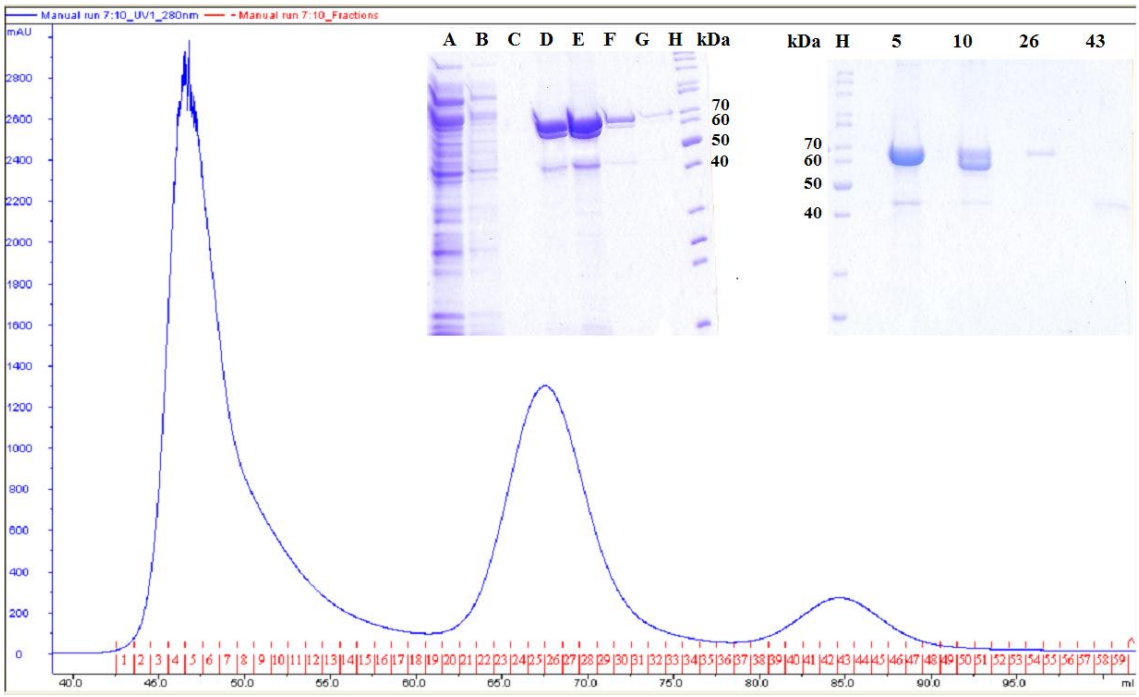
Tabela 5.13. Protokół oczyszczania rekombinowanego białka NodA-20-MBP.

etap	stosowane materiały	procedura
liza	roztwór 1: 20 mM Tris pH 7.5 500 mM NaCl 10% glicerol 10 mM EDTA PMSF (do stężenia 1 mM) lizozym (do stężenia 1 mg/mL)	- rozpuszczenie osadu komórkowego w roztworze 1 (ok. 40 mL / 4 g osadu) - dodanie PMSF i lizozymu - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - 1 godz. delikatnego mieszania w 20ºC - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - wirowanie: 20000 rpm, 4ºC, 30 min. - odrzucenie osadu
wysalanie i odsalanie	$(NH_4)_2SO_4$ w postaci rozdrobnionych kryształków, 56 g na 100 mL supernatantu cały etap prowadzony w temperaturze 4ºC roztwór 2: 20 mM Tris pH 7.5 500 mM NaCl 10% glicerol	- umieszczenie supernatantu w naczyniu pozwalającym na ciągłe delikatne mieszanie - dosypywanie do supernatantu kolejnych małych porcji siarczanu amonu aż do osiągnięcia 100% nasycenia (ok. 4 godz.) - wirowanie: 17000 rpm, 4ºC, 30 min. - odrzucenie supernatantu - rozpuszczanie osadu w roztworze 2 (ok 10 mL) - kilkukrotna dializa do roztworu 2 prowadzona w małych objętościach pozwalających na powolne obniżanie stężenia siarczanu amonu w preparacie białkowym
chromatografia powinowactwa na kolumnie amylozowej metodą grawitacyjną	złoże: Amylose Resin (NEB) roztwór 3: roztwór 2 + 100 mM maltoza cały etap prowadzony w temperaturze 4ºC	- zmieszanie supernatantu ze złożem wypełniającym kolumnę (ok. 10 mL złoża / 40 mL supernatantu) - inkubacja 15 - 30 min. - odrzucenie eluatu zawierającego frakcję białek niezwiązanych do złoża - przemywanie roztworem 2 (8 - 12 x 30 mL) - wymywanie białka NodA-20-MBP roztworem 3 (ok. 20 mL) - zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (Millipore)
chromatografia żelowa przy użyciu FPLC	kolumna GE HiLoad 16/60 Superdex 200 (GE-Healthcare)	- nałożenie preparatu na kolumnę - rozdział przy użyciu roztworu 2 - połączenie frakcji zawierających natywne białko - zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (Millipore)

Optymalizacja opracowanego protokołu dotyczyła zarówno zawartych w nim etapów jak i składników używanych roztworów. Przede wszystkim próbowałam oddzielić NodA od MBP za pomocą proteazy - czynnika Xa (Novagen) i zgodnie z protokołem dostarczonym wraz z enzymem. Skutkowało to jednak całkowitym strącaniem się NodA, podczas gdy białko MBP pozostawało w roztworze. Modyfikacja polegająca na dodaniu etapu chromatografii jonowymiennej powodowała znaczne obniżenie wydajności. Z kolei pominięcie czasochłonnego etapu wysalania i odsalania

skutkowało pogorszeniem jakości preparatu NodA-20-MBP, który w całości przechodził do frakcji agregatów (pierwszy pik chromatografii żelowej).

Rysunek 5.9. Przykładowy wynik oczyszczania białka NodA-20-MBP. Ścieżki na żelach SDS-PAGE: A-G- etap chromatografii powinowactwa: A- frakcja białek nie związanych do złoża; B, C- przemywanie złoża; D, E, F, G- eluat; H- marker masy; 5, 10, 26, 43- etap chromatografii żelowej, kolejne ścieżki to białka z frakcji FPLC oznaczone tymi samymi numerami, co na chromatogramie. Za pomocą spektrometrii mass zidentyfikowano białka zawarte w konkretnych prążkach wyciętych z żelu SDS-PAGE. Ścieżka/Frakcja 5 zawiera NodA-20-MBP (prążek ok. 66 kDa) i białko MBP z *E. coli* (prążek ok. 42 kDa), Ścieżka/Frakcja 10 zawiera dodatkowo białko GroEL hsp60 z *E. coli* (prążek ok. 60 kDa). Ścieżka/Frakcja 26 zawiera jedynie NodA-20-MBP (prążek ok. 66 kDa). Ścieżka/Frakcja 43 zawiera jedynie białko MBP z *E. coli* (słaby prążek ok. 42 kDa).



Próbując zwiększyć ilość otrzymywanego czystego preparatu NodA-20-MBP (tj. zwiększyć wypływ białek w drugim piku chromatografii żelowej) przetestowałam również wpływ składu używanych roztworów. Zmiany wprowadzałam na etapie odsalania, jednak ich efekt widoczny był dopiero na chromatogramach FPLC. Zarówno zwiększenie (do 1 M) jak i zmniejszenie (do 0.1 M) stężenia NaCl, jak i całkowite usunięcie z roztworu glicerolu, skutkowały jednak wyłącznie agregacją białek (zanik drugiego piku).

Pomiary DLS oraz krystalizacja białka

Uzyskiwany przeze mnie czysty preparat NodA-20-MBP pochodzący z drugiego piku chromatografii żelowej, w roztworze zawierającym 20 mM Tris pH 7.5, 500 mM

NaCl i 5% glicerol, zatężałam do stężenia 6 mg/mL (maksymalne stężenie nie powodujące strącania). Następnie poddawałam preparat krystalizacji oraz testom dynamicznego rozpraszania światła (DLS) na analizatorze Zetasizer Nano (Malvern Instruments Ltd.) (Rysunek 5.10).

Analizator Zetasizer wykazał, że preparat w zasadzie nie zawiera agregatów, ponieważ cząsteczki o znacznych wielkościach (pik nr 2 na Rysunku 5.10) stanowią zaledwie 0.1% masy próbki. Jednocześnie znaczna polidyspersja (%Pd 31.6) frakcji stanowiącego 99.9% masy próbki (pik nr 1 na Rysunku 5.10) świadczy o niejednorodności tego preparatu. W przypadku białka NodA-20-MBP potencjalne przyczyny takiego stanu mogą być różne, np. elastyczna konstrukcja pozwalająca na różnorodne konformacje NodA i MBP, czy nieokreślona struktura czwartorzędowa i możliwość jednoczesnego występowania di- tri- czy tetra-merów. Z tych samych powodów również wyznaczony promień hydrodynamiczny i wyliczona z niego masa białka mogą być obarczone sporym błędem.

Rysunek 5.10. Przykładowy wynik pomiaru DLS preparatu NodA-20-MBP zawierającego białka pochodzące z drugiego piku chromatografii żelowej w roztworze składającym się z: 20 mM Tris pH 7.5, 500 mM NaCl, 5% glicerol.

ASGLIC_frakcja26 3

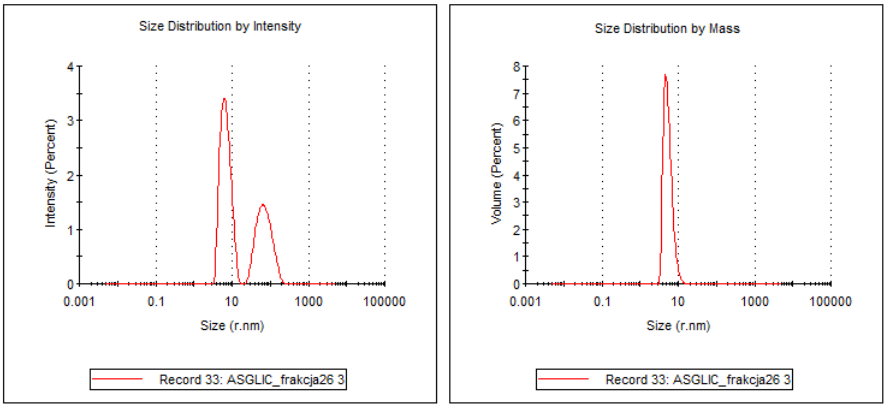
Record Number:	33	**File Name:**	AX_po_GF_.dts
Temperature (°C):		**Meas Date & Time:**	28 January 2011 14:58:39
Derived Count Rate (kcps):	431.4	**Solvent:**	buffer_bejger_Tris20_NaCl500
Intercept:	0.527	**Measurement quality:**	Good

Hydrodynamic Radius

Z-Average (± SD) (r.nm):	9.142 ± 7.205	**Estimated MW (± SD) (KDa):**	597 ± 470.3
Polydispersity Index:	0.621	**%Polydispersity:**	78.8
		Sample Polydispersity:	Polydisperse

Distribution Results

	Mode ± SD (nm)	%Pd	Est. MW (KDa) (Mode ± SD)*	% Intensity	% Mass	Peak Polydispersity
Peak 1:	6.156 ± 2.219	31.6	236.5 ± 101.7	61.5	99.9	Polydisperse
Peak 2:	64.97 ± 34.53	46.1	5.87e+4 ± 3.78e+4	38.5	0.1	Polydisperse
Peak 3:	0.000 ± 0.000	0	0.0 ± 0.0	0.0	0.0	



*The molecular weight reported here is only an estimate, calculated using an empirical mass vs. size calibration curve.

Białko NodA-20-MBP poddawałam krystalizacji korzystając z komercyjnych zestawów do krystalizacji firm Hampton i MDL, jak również z samodzielnie przygotowanego zestawu (tzw. MBP-screen-24), który zaprojektowałam opierając się na wyszukanych w literaturze informacjach dotyczących warunków krystalizacji białka MBP (Tabela 5.14). Początkowo nastawiałam ręcznie krople wiszące, a później również siedzące, korzystając z robota krystalizacyjnego Gryphon (Art Robbins Instruments), w rozmaitych proporcjach preparatu do rezerwuaru. Niestety w doświadczeniach tych nie uzyskałam żadnych kryształów białka.

Tabela 5.14. Zestaw MBP-screeen-24 do krystalizacji białka MBP metodą kropli wiszącej zawierający 24 roztwory.

16% PEG400 200 mM MES pH 5.0	16% PEG400 200 mM cytrynian sodu pH 4.0	16% PEG400 200 mM kakodylan sodu pH 5.0	16% PEG400 200 mM octan sodu pH 4.0
16% PEG400 200 mM MES pH 5.3	16% PEG400 200 mM cytrynian sodu pH 4.5	16% PEG400 200 mM kakodylan sodu pH 5.5	16% PEG400 200 mM octan sodu pH 4.3
16% PEG400 200 mM MES pH 5.6	16% PEG400 200 mM cytrynian sodu pH 5.0	16% PEG400 200 mM kakodylan sodu pH 6.0	16% PEG400 200 mM octan sodu pH 4.6
16% PEG400 200 mM MES pH 6.0	16% PEG400 200 mM cytrynian sodu pH 5.5	16% PEG400 200 mM kakodylan sodu pH 6.5	16% PEG400 200 mM octan sodu pH 5.0
16% PEG400 200 mM MES pH 6.3	16% PEG400 200 mM cytrynian sodu pH 6.0	16% PEG400 200 mM kakodylan sodu pH 7.0	16% PEG400 200 mM octan sodu pH 5.3
16% PEG400 200 mM MES pH 6.6	16% PEG400 200 mM cytrynian sodu pH 6.5	16% PEG400 200 mM kakodylan sodu pH 7.5	16% PEG400 200 mM octan sodu pH 5.6

5.2.4.4. NodA z *Bradyrhizobium* sp. WM9 z białkiem fuzyjnym MBP (wersja 2.)

Gdy korzystałam z wektora pMB-wm9-NodA, białko NodA z *Bradyrhizobium* sp. WM9 ulegało ekspresji w fuzji z białkiem MBP (63.4 kDa). Tym razem jednak oba białka połączone były krótkim pięcioaminokwasowym łącznikiem, zmniejszającym elastyczność tandemu białek oraz nie posiadającym miejsca cięcia proteazą.

Produkcja białka

Ekspresja NodA sztywno związanego z MBP (białko oznaczone skrótem NodA-5-MBP) pozwoliła uzyskać rozpuszczalne białka, stanowiące ok. 20% całej nadprodukcji.

Rozpuszczalne białka przeznaczone do oczyszczania były produkowane zgodnie z protokołem zamieszczonym w Tabeli 5.15.

Tabela 5.15. Protokół produkcji rekombinowanego białka NodA-5-MBP.

etap	stosowane materiały	procedura
transformacja	wektor pMB-wm9-NodA komórki BL21 magic (MCSG) pożywka płynna SOC	- delikatne zmieszanie komórek z wektorem (ok. 5 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5min - zmieszanie komórek z SOC (250 µL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla na pożywce stałej	pożywka stała LB z ampicyliną (0.1 mg/mL) z kanamycyną (25 µg/mL)	- posiew na płytkę z pożywką stałą LB - inkubacja 37ºC, 16 godz.
hodowla na pożywce płynnej	pożywka płynna TB z glukozą (0.2% w/v) z ampicyliną (0.1 mg/mL) z kanamycyną (25 µg/mL)	- przeniesienie komórek z płytki do pożywki płynnej TB (250 mL) - inkubacja 37ºC, 2 godz., 250 rpm - rozcieńczenie hodowli świeżą pożywką TB (do 2 L) - inkubacja 37ºC, 250 rpm, do OD_{600}=0.8 (ok 3 godz.)
indukcja białek opiekuńczych		- inkubacja 42ºC, 250 rpm, 0.5 godz. - obniżenie temperatury do 18ºC
ekspresja	IPTG (do stężenia 0.3 mM)	- dodanie IPTG - inkubacja 18ºC, 16 godz., 250 rpm
zabezpieczenie komórek		- wirowanie 5000 rmp, 10ºC, 30 min. - zamrożenie komórek w -80ºC

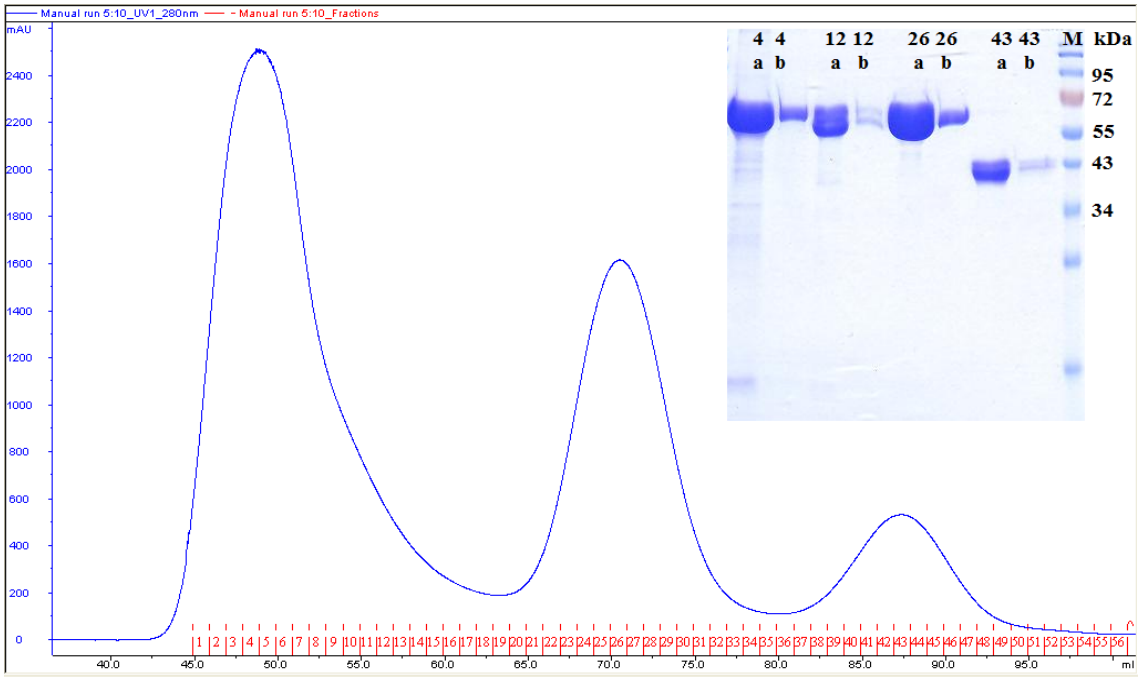
Oczyszczanie białka

Rozpoczynając prace z NodA-5-MBP korzystałam z protokołu przygotowanego dla NodA-20-MBP i uzyskałam podobny wynik, czyli dwa piki na chromatogramie FPLC zawierające NodA. Protokół ten zmodyfikowałam (Tabela 5.16), choć podobnie jak w przypadku NodA-20-MBP, nie udało mi się zapobiec stratom białka następującym podczas oczyszczania. Przede wszystkim usunęłam z protokołu pracochłonny i czasochłonny etap wysalania i odsalania, który, jak się okazało, w przypadku białka NodA-5-MBP nie był niezbędny. Do procedury wprowadziłam natomiast dodatkowy etap: zmianę składu roztworu białka, bezpośrednio po którym następowała chromatografia żelowa przy użyciu FPLC. Chromatogram FPLC pozwalał mi weryfikować jakość białka NodA-5-MBP w każdym nowym roztworze, ujawniając ilościowe proporcje rozdzielonych białek: w natywnej konformacji (drugi pik) i zagregowanych (pierwszy pik) (Rysunek 5.11). Przetestowałam w ten sposób kilka odczynników (Tabe-

Tabela 5.16. Protokół oczyszczania rekombinowanego białka NodA-5-MBP.

etap	stosowane materiały	procedura
liza	roztwór 1: 20 mM Tris pH 7.5 500 mM NaCl 10% glicerol PMSF (do stężenia 1 mM) lizozym (do stężenia 1 mg/mL)	- rozpuszczenie osadu komórkowego w roztworze 1 (ok. 40 mL / 4 g osadu) - dodanie PMSF i lizozymu - 0.5 - 1 godz. delikatnego mieszania w 20ºC - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - 0.5 - 1 godz. delikatnego mieszania w 20ºC - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - wirowanie: 20000 rpm, 4ºC, 30 min. - odrzucenie osadu
chromatografia powinowactwa na kolumnie amylozowej metodą grawitacyjną	złoże: Amylose Resins (NEB) roztwór 2: roztwór 1 + 100 mM maltoza cały etap prowadzony w temperaturze 4ºC	- zmieszanie supernatantu ze złożem (ok. 10 mL złoża / 40 mL supernatantu) - inkubacja 15 - 30 min. - odrzucenie eluatu zawierającego frakcję białek niezwiązanych do złoża - przemywanie roztworem 1 (8 - 12 x 30 mL) - wymywanie białka roztworem 2 (ok. 20 mL)
zmiana składu roztworu białka	roztwór 3: (szczegóły w Tabeli 5.17)	- dializa do roztworu 3 - zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (Millipore)
chromatografia żelowa przy użyciu FPLC	kolumna GE HiLoad 16/60 Superdex 200 (GE-Healthcare)	- nałożenie preparatu na kolumnę - rozdział przy użyciu roztworu 3 - połączenie frakcji zawierających natywne białko - zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (Millipore)

Rysunek 5.11. Przykładowy wynik chromatografii żelowej dla białka NodA-5-MBP przedstawiony na chromatogramie FPLC i żelu SDS-PAGE. Ścieżki na żelu SDS-PAGE: M- marker masy; a- nałożona większa ilość próbki, b- nałożona mniejsza ilość próbki; 4, 12, 26, 43- białka z frakcji FPLC oznaczone tymi samymi numerami, co na chromatogramie. Za pomocą spektrometrii mass zidentyfikowano białka zawarte w konkretnych prążkach wyciętych z żelu SDS-PAGE. Ścieżka/Frakcja 4 zawiera NodA-5-MBP (prążek ok. 66 kDa) i inne białka z *E. coli*, Ścieżka/Frakcja 12 zawiera NodA-5-MBP (górny prążek) i białko GroEL hsp60 z *E. coli* (dolny prążek). Ścieżka/Frakcja 26 zawiera jedynie NodA-5-MBP (prążek ok. 66 kDa). Ścieżka/Frakcja 43 zawiera jedynie białko MBP z *E. coli* (prążek ok. 42 kDa).



la 5.17), które jednak nie wpłynęły na poprawę stabilności białka NodA-5-MBP. Nowy protokół pozwalał mi dość szybko i ze stałą wydajnością oczyszczać białko (pochodzące wyłącznie z drugiego piku chromatografii żelowej) w różnych roztworach (Tabela 5.17) i w ten sposób przygotowywać różne preparaty NodA-5-MBP.

Tabela 5.17. Różne wersje roztworu 3, używanego podczas oczyszczania FPLC oraz obecnego w kolejnych krystalizowanych preparatach białka NodA-5-MBP.

20 mM Tris pH 7.5 500 mM NaCl 5% glicerol	20 mM Tris pH 7.5 500 mM NaCl	20 mM Tris pH 7.5 500 mM NaCl 10 mM β-merkaptoetanol	20 mM Tris pH 7.5 500 mM NaCl 5 mM maltoza

Pomiary DLS oraz krystalizacja białka

Uzyskiwane przeze mnie preparaty NodA-5-MBP zatężone do ok. 7 mg/mL poddawałam krystalizacji oraz testom DLS (Rysunek 5.12).

Rysunek 5.12. Przykładowy wynik pomiaru DLS preparatu NodA-5-MBP w roztworze zawierającym jedynie 20 mM Tris pH 7.5 i 500 mM NaCl.

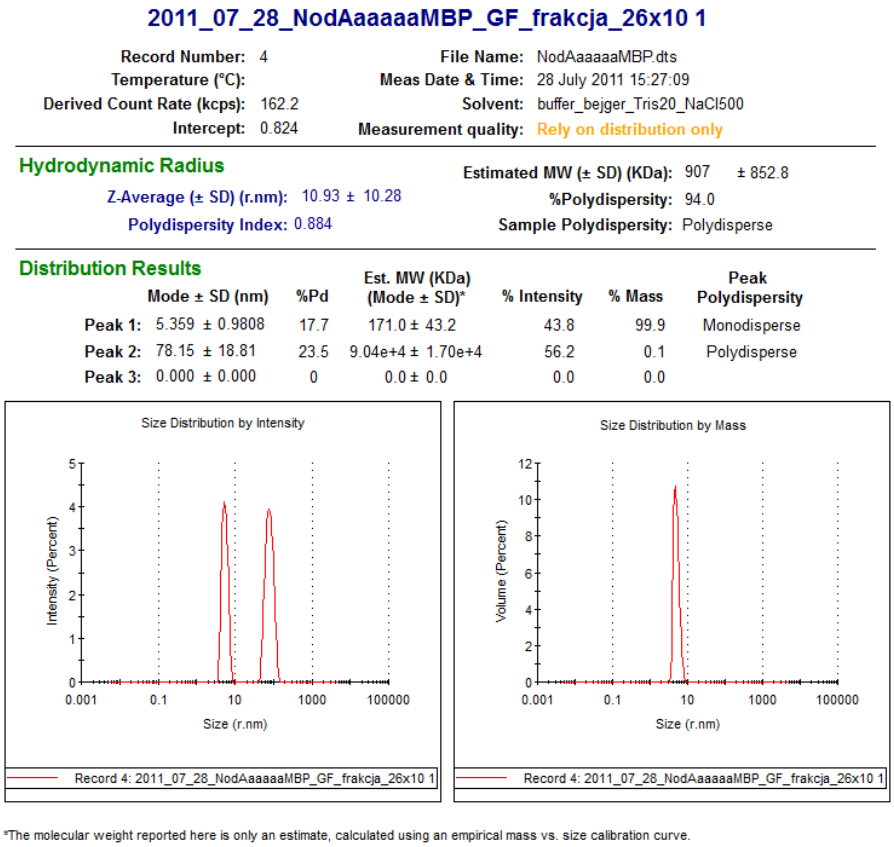


Tabela 5.18. Zestaw MBP-screeen-96 do krystalizacji białka MBP metodą kropli siedzącej przy użyciu robota krystalizacyjnego zawierający 96 roztworów.

200 mM (pH 6.4) MES 10% PEG400	200 mM (pH 6.2) MES 10% PEG400	200 mM (pH 6.0) MES 10% PEG400	200 mM (pH 5.8) MES 10 % PEG400	200 mM (pH 5.6) MES 10 % PEG400	200 mM (pH 5.4) MES 10 % PEG400	200 mM (pH 5.2) MES 10 % PEG400	200 mM (pH 5.0) MES 10% PEG400
200 mM (pH 6.4) MES 20% PEG400	200 mM (pH 6.2) MES 20% PEG400	200 mM (pH 6.0) MES 20% PEG400	200 mM (pH 5.8) MES 20% PEG400	200 mM (pH 5.6) MES 20 % PEG400	200 mM (pH 5.4) MES 20 % PEG400	200 mM (pH 5.2) MES 20 % PEG400	200 mM (pH 5.0) MES 20% PEG400
200 mM (pH 6.4) MES 30% PEG400	200 mM (pH 6.2) MES 30% PEG400	200 mM (pH 6.0) MES 30% PEG400	200 mM (pH 5.8) MES 30% PEG400	200 mM (pH 5.6) MES 30 % PEG400	200 mM (pH 5.4) MES 30 % PEG400	200 mM (pH 5.2) MES 30 % PEG400	200 mM (pH 5.0) MES 30% PEG400
200 mM (pH 6.5) cytrynian sodu 10% PEG400	200 mM (pH 6.2) cytrynian sodu 10% PEG400	200 mM (pH 5.9) cytrynian sodu 10% PEG400	200 mM (pH 5.6) cytrynian sodu 10% PEG400	200 mM (pH 5.3) cytrynian sodu 10 % PEG400	200 mM (pH 5.0) cytrynian sodu 10 % PEG400	200 mM (pH 4.7) cytrynian sodu 10 % PEG400	200 mM (pH 4.4) cytrynian sodu 10% PEG400
200 mM (pH 6.5) cytrynian sodu 20% PEG400	200 mM (pH 6.2) cytrynian sodu 20% PEG400	200 mM (pH 5.9) cytrynian sodu 20% PEG400	200 mM (pH 5.6) cytrynian sodu 20% PEG400	200 mM (pH 5.3) cytrynian sodu 20 % PEG400	200 mM (pH 5.0) cytrynian sodu 20 % PEG400	200 mM (pH 4.7) cytrynian sodu 20 % PEG400	200 mM (pH 4.4) cytrynian sodu 20% PEG400
200 mM (pH 6.5) cytrynian sodu 30% PEG400	200 mM (pH 6.2) cytrynian sodu 30% PEG400	200 mM (pH 5.9) cytrynian sodu 30% PEG400	200 mM (pH 5.6) cytrynian sodu 30% PEG400	200 mM (pH 5.3) cytrynian sodu 30 % PEG400	200 mM (pH 5.0) cytrynian sodu 30 % PEG400	200 mM (pH 4.7) cytrynian sodu 30 % PEG400	200 mM (pH 4.4) cytrynian sodu 30% PEG400
200 mM (pH 6.4) kakodylan sodu 10% PEG400	200 mM (pH 6.2) kakodylan sodu 10% PEG400	200 mM (pH 6.0) kakodylan sodu 10% PEG400	200 mM (pH 5.8) kakodylan sodu 10% PEG400	200 mM (pH 5.6) kakodylan sodu 10 % PEG400	200 mM (pH 5.4) kakodylan sodu 10 % PEG400	200 mM (pH 5.2) kakodylan sodu 10 % PEG400	200 mM (pH 5.0) kakodylan sodu 10% PEG400
200 mM (pH 6.4) kakodylan sodu 20% PEG400	200 mM (pH 6.2) kakodylan sodu 20% PEG400	200 mM (pH 6.0) kakodylan sodu 20% PEG400	200 mM (pH 5.8) kakodylan sodu 20% PEG400	200 mM (pH 5.6) kakodylan sodu 20 % PEG400	200 mM (pH 5.4) kakodylan sodu 20 % PEG400	200 mM (pH 5.2) kakodylan sodu 20 % PEG400	200 mM (pH 5.0) kakodylan sodu 20 % PEG400
200 mM (pH 6.4) kakodylan sodu 30% PEG400	200 mM (pH 6.2) kakodylan sodu 30% PEG400	200 mM (pH 6.0) kakodylan sodu 30% PEG400	200 mM (pH 5.8) kakodylan sodu 30% PEG400	200 mM (pH 5.6) kakodylan sodu 30 % PEG400	200 mM (pH 5.4) kakodylan sodu 30 % PEG400	200 mM (pH 5.2) kakodylan sodu 30 % PEG400	200 mM (pH 5.0) kakodylan sodu 30 % PEG400
200 mM (pH 5.6) octan sodu 10% PEG400	200 mM (pH 5.4) octan sodu 10% PEG400	200 mM (pH 5.2) octan sodu 10% PEG400	200 mM (pH 5.0) octan sodu 10% PEG400	200 mM (pH 4.8) octan sodu 10 % PEG400	200 mM (pH 4.6) octan sodu 10 % PEG400	200 mM (pH 4.4) octan sodu 10 % PEG400	200 mM (pH 4.2) octan sodu 10 % PEG400
200 mM (pH 5.6) octan sodu 20% PEG400	200 mM (pH 5.4) octan sodu 20% PEG400	200 mM (pH 5.2) octan sodu 20% PEG400	200 mM (pH 5.0) octan sodu 20% PEG400	200 mM (pH 4.8) octan sodu 20 % PEG400	200 mM (pH 4.6) octan sodu 20 % PEG400	200 mM (pH 4.4) octan sodu 20 % PEG400	200 mM (pH 4.2) octan sodu 20 % PEG400
200 mM (pH 5.6) octan sodu 30% PEG400	200 mM (pH 5.4) octan sodu 30% PEG400	200 mM (pH 5.2) octan sodu 30% PEG400	200 mM (pH 5.0) octan sodu 30% PEG400	200 mM (pH 4.8) octan sodu 30 % PEG400	200 mM (pH 4.6) octan sodu 30 % PEG400	200 mM (pH 4.4) octan sodu 30 % PEG400	200 mM (pH 4.2) octan sodu 30 % PEG400

Analizator Zetasizer Nano wykazał, że wszystkie badane próbki białka NodA-5-MBP, niezależnie od składu roztworów (Tabela 5.17) były bardzo podobne. Preparaty te w całości (w 99.9% masy) stanowiły homogenne roztwory (%Pd 17.7), nie zawierające agregatów. Obniżenie polidyspersji białka NodA-5-MBP w porównaniu do NodA-20-MBP mogło być spowodowane usztywnieniem konstrukcji białko-łącznik-białko. Uzyskiwane promienie hydrodynamiczne były nieznacznie rozbieżne (5.2 - 6.2 nm) co było głównie efektem błędu pomiarowego wynikającego z różnych lepkości preparatów. Oszacowana masa białka sugeruje że Noda-5-MBP jest trimerem, nie jest to jednak pewne, ponieważ została ona wyliczona z promienia dla obiektu idealnie sferycznego co, szczególnie dla białek fuzyjnych, jest dalekie do prawdy.

Każdy z preparatów NodA-5-MBP poddawałam krystalizacji przesiewowej używając robota krystalizacyjnego oraz korzystając z komercyjnych zestawów do krystalizacji firm Hampton i MDL, jak również z własnego (zawierającego 96 roztworów) zestawu do krystalizacji MBP, tzw. MBP-screen-96 (Tabela 5.18). Niestety i tym razem nie udało mi się uzyskać monokryształów białka.

5.2.4.5. NodB z *Bradyrhizobium* sp. WM9 z białkiem fuzyjnym MBP

Z wektora pMB-wm9-NodB białko NodB z *Bradyrhizobium* sp. WM9 ulegało ekspresji w fuzji z białkiem MBP (63.5 kDa). Oba białka złączone były pięcioaminokwasowym łącznikiem uniemożliwiającym ich oddzielenie (brak miejsca cięcia) oraz usztywniającym konstrukcję fuzji. Sekwencja białka NodB z *Bradyrhizobium* sp. WM9 zawiera 7 reszt cysteiny, więc podczas oczyszczania stosowałam czynniki redukujące.

Produkcja białka

Ekspresja na dużą skalę NodB sztywno związanego z MBP (białko oznaczone skrótem NodB-5-MBP) pozwoliła uzyskać małą pulę rozpuszczalnych białek, stanowiących maksymalnie 10% całej nadprodukcji. Rozpuszczalne białko przeznaczone do oczyszczania było produkowane zgodnie z protokołem identycznym jak protokół produkcji białka NodA-5-MBP (Tabela 5.15).

Oczyszczanie białka

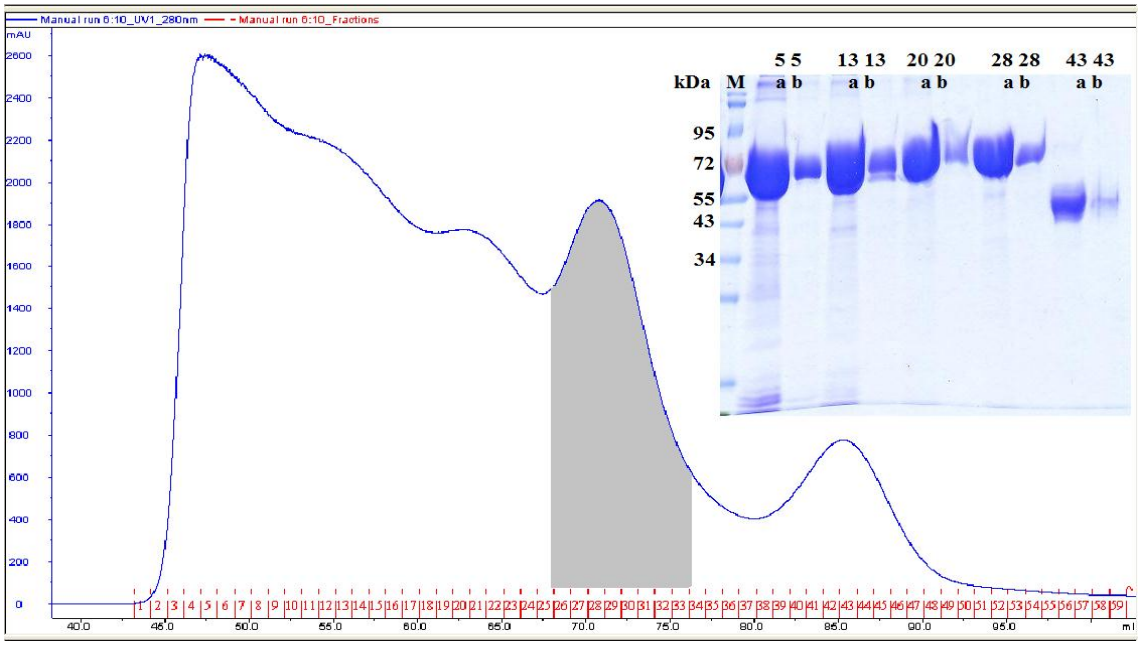
Białko NodB-5-MBP było przeze mnie oczyszczane według protokołu wzorowanego na NodA-5-MBP (Tabela 5.19). Tak jak poprzednio, etap chromatografii żelowej wykorzystywałam do weryfikacji kolejno wprowadzanych zmian w składzie roztworu 3, mających na celu poprawę wydajności procedury.

Tabela 5.19. Protokół oczyszczania rekombinowanego białka NodB-5-MBP.

etap	stosowane materiały	procedura
liza	roztwór 1: 20 mM Tris pH 7.5 200 mM NaCl 10% glicerol 10 mM β-merkaptoetanol PMSF (do stężenia 1 mM) lizozym (do stężenia 1 mg/mL)	- rozpuszczenie osadu komórkowego w roztworze 1 (ok. 40 mL / 4 g osadu) - dodanie PMSF i lizozymu - 0.5 - 1 godz. delikatnego mieszania w 20ºC - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - 0.5 - 1 godz. delikatnego mieszania w 20ºC - sonikacja (6 - 10 x 10 - 15 sek., moc 80%) - wirowanie: 20000 rpm, 4ºC, 30 min. - odrzucenie osadu
chromatografia powinowactwa na kolumnie amylozowej metodą grawitacyjną	złoże: Amylose Resins (NEB) roztwór 2: roztwór 1 + 100 mM maltoza cały etap prowadzony w temperaturze 4ºC	- zmieszanie supernatantu ze złożem (ok. 10 mL złoża / 40 mL supernatantu) - inkubacja 15 - 30 min. - odrzucenie eluatu zawierającego frakcję białek niezwiązanych do złoża - przemywanie roztworem 1 (8 - 12 x 30 mL) - wymywanie białka roztworem 2 (ok. 20 mL)
zmiana składu roztworu białka	roztwór 3: (składniki wymieniłam w tekście)	- dializa do roztworu 3 - zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (Millipore)
chromatografia żelowa przy użyciu FPLC	kolumna GE HiLoad 16/60 Superdex 200 (GE-Healthcare)	- nałożenie preparatu na kolumnę - rozdział przy użyciu roztworu 3 - zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (Millipore)

Chromatografię żelową prowadziłam między innymi w trzech stężeniach NaCl (100 mM, 200 mM, 500 mM), w obecności β-merkaptoetanolu (10 mM), TCEP (3 mM), maltozy (10 mM) i glicerolu (10%). Białko silnie agregowało w 500 mM NaCl (wypływ tylko w piku pierwszym), a podczas dializy, która całkowicie usunęła sól, strąciło się zupełnie uniemożliwiając wykonanie chromatografii. Pozostałe stężenia i odczynniki nie miały większego wpływu na rozdział FPLC, a chromatogramy za każdym razem zawierały kilka zachodzących na siebie pików zawierających to samo białko, NodB-5-MBP (Rysunek 5.13). Próby wydzielenia białek pochodzących z jednego piku, zatężenie ich i ponowna chromatografia żelowa dawały praktycznie identyczne chromatogramy, z kilkoma pikami zawierającymi NodB-5-MBP.

Rysunek 5.13. Przykładowy wynik chromatografii żelowej dla białka NodB-5-MBP na chromatogramie FPLC i żelu SDS-PAGE. Ścieżki na żelu: M- marker masy, a- nałożona większa ilość próbki, b- nałożona mniejsza ilość próbki; 5, 13, 20, 28, 43- białka z frakcji FPLC oznaczone tymi samymi numerami, co na chromatogramie. NodB-5-MBP (prążek ok. 64 kDa) występuje we wszystkich frakcjach od 1 do 33. Ścieżka/Frakcja 13 zawiera także białko GroEL hsp60 z *E. coli* (dolny prążek). Ścieżka/Frakcja 43 zawiera jedynie białko MBP z *E. coli* (prążek ok. 42 kDa). Obszar zaznaczony na szaro określa frakcje używane do dalszych prac (ponowna chromatografia żelowa, DLS, krystalizacja).

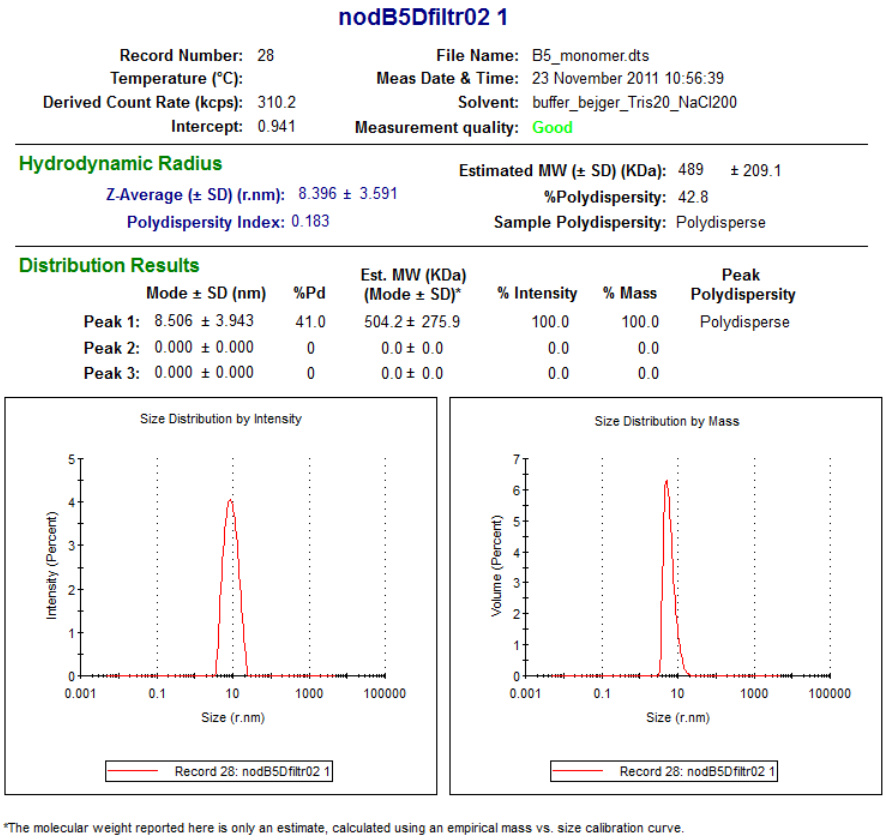


Pomiary DLS oraz krystalizacja białka

Białko NodB-5-MBP pochodzące z tych frakcji chromatografii żelowej, które nie zawierały innych białek (pomiędzy 70 i 75 mL objętości wypływającej z kolumny) mogłam zatężać nawet do 15 mg/mL. Jednak pomiary DLS takich preparatów ukazywały ogromny rozrzut wielkości cząsteczek zawartych w badanych próbkach (Rysunek 5.14). Preparaty NodB-5-MBP najczęściej nie zawierały agregatów, cechowała je natomiast znaczna polidyspersja dochodząca nawet do 66%, natomiast wartość promienia hydrodynamicznego wahała się w przedziale 3.7 - 11.8 nm, co odpowiadało zakresowi masy 70 - 1000 kDa. Oczywiście różne wyniki otrzymywałam dla różnych próbek (o różnym składzie odczynników, stężeniach, temperaturze, czasie przechowywania), jednak pomiary te nie były powtarzalne i nie mogłam zauważyć żadnej konkretnej i użytecznej zależności.

Preparaty NodB-5-MBP poddawałam krystalizacji przesiewowej posługując się robotem krystalizacyjnym. Używałam komercyjnych zestawów do krystalizacji firm Hampton i MDL oraz zestawu MBP-screen-96 (Tabela 5.18). Niestety i tym razem nie udało mi się uzyskać kryształów badanego białka.

Rysunek 5.14. Przykładowy wynik pomiaru DLS preparatu NodB-5-MBP.



5.2.4.6. NoeI z *Bradyrhizobium japonicum* z białkiem fuzyjnym MBP

Z wektora pMB-jap-NoeI białko NoeI z *Bradyrhizobium japonicum* USDA110 ulegało ekspresji w fuzji z białkiem MBP (64.4 kDa). Oba białka złączone były pięcio-aminokwasowym łącznikiem uniemożliwiającym ich rozdzielenie (brak miejsca cięcia) oraz usztywniającym konstrukcję fuzji. Sekwencja białka NoeI z *B. japonicum* zawiera 4 reszty Cys, w związku z czym podczas oczyszczania stosowałam czynniki redukujące.

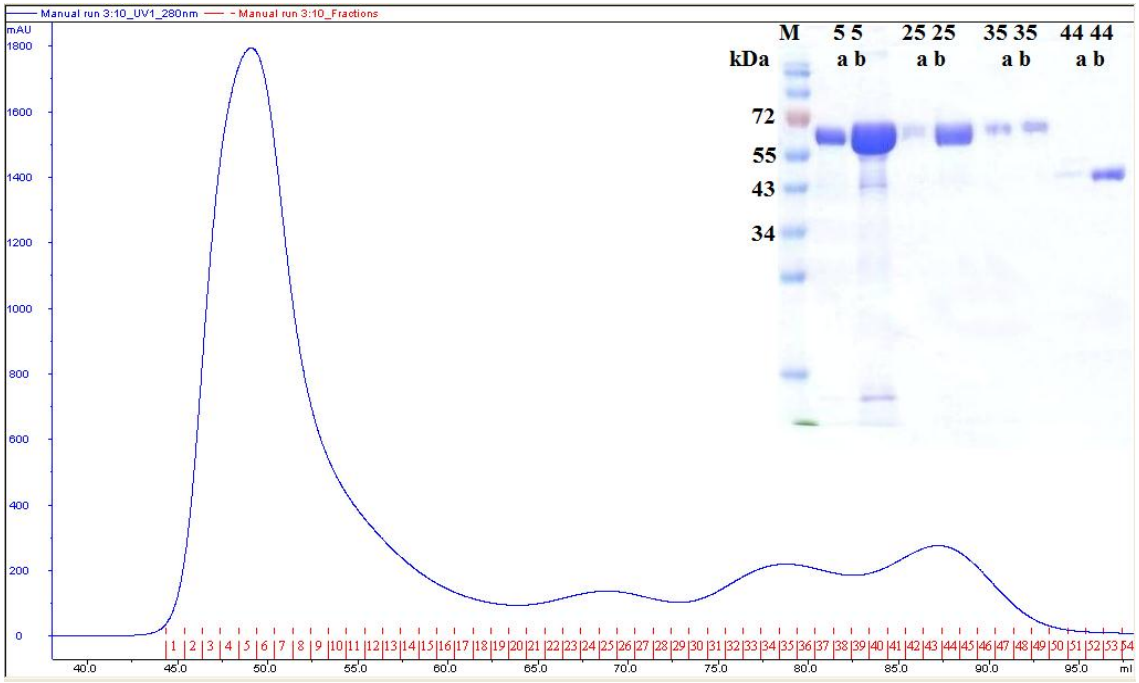
Produkcja białka

Ekspresja na dużą skalę NoeI sztywno związanego z MBP (białko NoeI-5-MBP) pozwoliła uzyskać małą pulę rozpuszczalnych białek, stanowiących maksymalnie 10% całej nadprodukcji. Rozpuszczalne białko przeznaczone do oczyszczania było produkowane zgodnie z protokołem identycznym jak protokół produkcji białka NodA-5-MBP (Tabela 5.15).

Oczyszczanie białka

Białko NoeI-5-MBP oczyszczałam podobnie jak dwa pozostałe białka nodulacj o tej samej konstrukcji. Korzystałam z tego samego protokołu co podczas prac z NodB-5-MBP (Tabela 5.19) i tak samo poszukiwałam optymalnego składu roztworu nr 3. Podobnie jak poprzednio przetestowałam wpływ różnych stężeń NaCl (200 mM i 500 mM), obecności β-merkaptoetanolu (10 mM) i glicerolu (10%). Za każdym razem jednak otrzymywałam ten sam wynik (Rysunek 5.15). Na końcu procedury, znaczna większość oczyszczanych rozpuszczalnych makrocząsteczek NoeI-5-MBP stanowiła wyłącznie agregaty i wypływała w chromatografii żelowej tuż po objętości zerowej kolumny. Pozostała niezagregowana pula białek NoeI-5-MBP, wypływająca jako dwa kolejne bardzo niskie piki, była bardzo mała i stanowiła mocno rozcieńczony preparat. Praca z takim preparatem byłaby nieopłacalna i miała małe szanse powodzenia. Stąd, w tym przypadku w ogóle nie przystąpiłam do krystalizacji białka.

Rysunek. 5.15. Przykładowy wynik chromatografii żelowej dla białka NoeI-5-MBP przedstawiony na chromatogramie FPLC i żelu SDS-PAGE. Ścieżki na żelu SDS-PAGE: M- marker masy, a- nałożona mniejsza ilość próbki, b- nałożona większa ilość próbki; 5, 25, 35, 44- białka z frakcji FPLC oznaczone tymi samymi numerami, co na chromatogramie. Za pomocą spektrometrii mass zidentyfikowano niektóre białka zawarte w konkretnych prążkach wyciętych z żelu SDS-PAGE. NoeI-5-MBP (prążek ok. 64 kDa) występuje we wszystkich frakcjach od 1 do 35. Ścieżka/Frakcja 44 zawiera jedynie białko MBP z *E. coli* (prążek ok. 42 kDa).



5.3. Dyskusja wyników

5.3.1. Analiza prób ekspresji, oczyszczania i krystalizacji białka NodA

Dla enzymu NodA przygotowałam cztery wersje białka rekombinowanego, ulegające ekspresji z czterech wektorów ekspresyjnych: NodA-wm9 z wektora pET-wm9-NodA, NodA-mel z wektora pET-mel-NodA, NodA-20-MBP z wektora pMAL-wm9-NodA i NodA-5-MBP z wektora pMB-wm9-NodA.

Pierwszy skonstruowany przeze mnie wektor (pET-wm9-NodA) pozwalał na produkcję białka NodA-wm9 jedynie w formie nierozpuszczalnej. W tym czasie inne białka nodulacji, nad którymi pracowałam, również ulegały ekspresji jedynie w ciałkach inkluzyjnych. Zdecydowałam się zatem spróbować oczyścić NodA-wm9 korzystając z technik umożliwiających zwijanie białek *in vitro* (*ang. refolding*). Jednak mimo starań, nie uzyskałam preparatu nadającego się do krystalizacji. Nierozpuszczalność NodA-wm9, zmusiła mnie do powrotu do etapu klonowania.

Postanowiłam zmienić źródło genu *nodA*, mając nadzieję, że różnice w sekwencjach homologów wpłyną na zmianę biofizycznej charakterystyki białka. Wybrałam *Sinorhizobium meliloti* 1021, ponieważ sekwencja NodA z tego szczepu znacznie różni się od sekwencji NodA z *Bradyrhizobium* sp. WM9 (identyczność 63%, podobieństwo 78%). Stosując nowy sposób (termiczną indukcję białek opiekuńczych) na poprawienie rozpuszczalności białek nadprodukowanych z wektora pET-mel-NodA, uzyskiwałam małą pulę białek rozpuszczalnych, dzięki której rozpoczęłam etap oczyszczania. Białko NodA-mel okazało się jednak bardzo niestabilne, szybko wytrącało się z roztworu oraz na filtrach. Jedyną szansę na uzyskanie preparatu (i przejście do etapu krystalizacji) dawała błyskawiczna procedura oczyszczania prowadzona w 4ºC. Ostatecznie białko NodA-mel i tak strącało się w kroplach krystalizacyjnych.

Ponownie zmodyfikowałam system ekspresyjny, tym razem zmieniając wykorzystywany wektor i skonstruowałam pMAL-wm9-NodA, który umożliwiał ekspresję białka nodulacji w fuzji z MBP. Nie wiadomo do końca jak działa białko MBP, jednak podejrzewa się, że promuje ono prawidłowe zwijanie się złączonych z nim białek, działając podobnie jak białka opiekuńcze, co w efekcie poprawia ich rozpuszczalność (**Fox *et al.*, 2001**). Białko NodA-20-MBP było znacznie stabilniejsze od poprzedniej wersji NodA, jednak dla zachowania rozpuszczalności musiało

pozostawać w fuzji z MBP. Pomiary DLS wykazały znaczną polidyspersje preparatu, świadczącą o braku homogenności strukturalnej, która przekłada się na mniejsze prawdopodobieństwo uzyskania kryształów, których rzeczywiście nie otrzymałam.

Aby zwiększyć szanse wykrystalizowania NodA, postanowiłam zmienić pMAL-wm9-NodA tak, aby uzyskiwana fuzja białek była mniej elastyczna. Zgodnie z danymi literaturowymi (**Smyth *et al.*, 2003**), aby umożliwić białkom w fuzji poprawne zwijanie się, należy pozostawić między nimi tylko krótki łącznik peptydowy. Fragment sekwencji wektora odpowiadający za ekspresję dwudziuestoaminokwasowego łącznika między NodA i MBP zamieniłam więc na sekwencję, której ekspresja pozwala uzyskać łącznik składający się jedynie z 5 reszt aminokwasowych. W ten sposób powstał wektor pMB-wm9-NodA. Usztywnienie konstrukcji NodA-5-MBP poprawiło stabilność białka. Dzięki temu mogłam uprościć procedurę oczyszczania oraz pozbyć się takich składników preparatu jak glicerol, który najczęściej hamuje krystalizację.

W przypadku białka NodA-5-MBP analizator DLS wykazał niską polidyspersję. Preparat ten poddawałam zakrojonym na szeroką skalę próbom krystalizacyjnym, korzystając z robota krystalizacyjnego. Ze względu na fakt, że 2/3 masy białka NodA-5-MBP stanowi MBP, starałam się otrzymać kryształy również stosując MBP-screen-96, czyli zestaw do krystalizacji zaprojektowany specjalnie dla białka MBP. Mimo wielu różnorodnych prób krystalizacyjnych, kryształu NodA nie udało mi się otrzymać.

5.3.2. Próby ekspresji, oczyszczania i krystalizacji pozostałych białek nodulacji

Dla białek NodB z *Bradyrhizobium* sp. WM9 i NoeI z *Bradyrhizobium japonicum* przygotowałam w sumie po trzy wektory ekspresyjne, każdy innego typu: pET-wm9-NodB, pMAL-wm9-NodB oraz pMB-wm9-NodB, a także pET-jap-NoeI, pMAL-jap-NoeI oraz pMB-jap-NoeI.

Białka NodB i NoeI ulegające ekspresji z wektorów typu pET i pMAL były nierozpuszczalne niezależnie od wprowadzanych modyfikacji fizyko-chemicznych warunków nadprodukcji. Dopiero ekspresja z wektora typu pMB w komórkach BL21 magic połączona z termiczną indukcją białek opiekuńczych oraz produkcją w 18ºC pozwoliła uzyskać pulę rozpuszczalnych białek NodB-5-MBP i NoeI-5-MBP.

Podczas oczyszczania białka NodB-5-MBP okazało się, że jest ono wysoce niehomogenne, niezależnie od zastosowanych warunków chemicznych. Szczególnie widoczne było to na chromatogramach FPLC i w wynikach z pomiarów DLS. Próby

oczyszczenia wybranej frakcji zawierającej danej wielkości białko kończyły się niepowodzeniem, podobnie jak nastawione krystalizacje. Z kolei próbując oczyścić preparat NoeI-5-MBP otrzymywałam za każdym razem białka, które stanowiły praktycznie same agregaty, w ogóle nie nadające się do krystalizacji.

Białka NodU i NolO, niezależnie od szczepu z którego pochodziły, typu wektora wykorzystanego do klonowania, rodzaju komórek gospodarza, w którym następowała ekspresja oraz warunków fizyko-chemicznych podczas nadprodukcji, zawsze były nierozpuszczalne.

5.3.3. Przydatność białek nodulacji do badań krystalograficznych

Głównymi problemami, jakie napotkałam w trakcie realizacji niniejszego projektu były nierozpuszczalność i niestabilność rekombinowanych białek nodulacji. Problemy te uniemożliwiły mi przejście do kolejnych etapów prac: nadekspresji, oczyszczania i krystalizacji.

Nie wiadomo dlaczego pochodzące z rizobiów białka nodulacji są nierozpuszczalne wewnątrz bakteryjnego gospodarza. Spośród dostępnych systemów ekspresyjnych, system oparty na *Escherichia coli* wydaje się najlepszym wyborem, ponieważ jest najbardziej zbliżony genetycznie. Istnieje spore prawdopodobieństwo, że białka nodulacji są toksyczne dla *E. coli*, dlatego organizm ten eliminuje je tworząc ciałka inkluzyjne. Możliwe też, że w naturalnym środowisku komórki riozobium białka te są w jakimś stopniu związane z błonami komórkowymi. Zawarta w literaturze charakterystyka bakteryjnych białek nodulacji jest dość skąpa. Nie wiadomo, jak białka te współdziałają produkując czynnik Nod, czy stanowią stały kompleks enzymatyczny, czy też działają raczej osobno, czy być może ich funkcjonowanie jest jeszcze bardziej skomplikowane.

Problemy nierozpuszczalności i niestabilności starałam się pokonać wieloma sposobami. Podczas ekspresji korzystałam z sekwencji różnych szczepów rizobiów (*Sinorhizobium meliloti* 1021, *Bradyrhizobium* sp. WM9, *Rhizobium* sp. NGR234 i *Bradyrhizobium japonicum* USDA110) oraz różnego typu wektorów (pET, pMAL, pMB). Produkcję prowadziłam w różnych komórkach ekspresyjnych (BL21 Star, NEB express, BL21 magic, OverExpress C41 i C43) oraz warunkach fizykochemicznych. Na etapie oczyszczania optymalizowałam procedury pod względem stosowanych etapów pracy oraz składników używanych roztworów.

Znalezione rozwiązania (takie jak indukcja białek opiekuńczych) sprawdzały się jednak tylko w niektórych przypadkach. Ciekawym przykładem może być zastosowanie fuzji białek. Obecność MBP (ok. 40 kDa) powodowała zwiększenie rozpuszczalności tylko u białek mających mniejszą masę (ok. 20 kDa, tj. NodA, NodB, NoeI), natomiast w przypadku NodU i NolO (ok. 70 kDa) nie dawała spodziewanego efektu. Jednocześnie z moich obserwacji wynika, że MBP nie jest gwarantem stabilności, gdyż nie udało mi się uzyskać homogennego preparatu dla białek NodB i NoeI. Z wszystkich wersji białek, nad którymi pracowałam w tym projekcie, jedynie preparat NodA-5-MBP nadawał się do krystalizacji. Choć nie uzyskałam monokryształów NodA, jestem przekonana, że kontynuacja moich badań nad tym białkiem może w przyszłości zaowocować udaną krystalizacją, a w konsekwencji strukturą krystaliczną. W konkluzji stwierdzam, że spośród przebadanych przeze mnie białek nodulacji, tylko NodA jest obiektem dającym nadzieję na pomyślny wynik badań krystalograficznych.

6. Asparaginaza z *Rhizobium etli* (AnsA)

6.1. Strategia postępowania z enzymem AnsA

Prace nad białkiem AnsA rozpoczęłam w oparciu o przygotowany wcześniej w naszej pracowni plazmidowy wektor ekspresyjny zawierający gen *ansA*, nazwany pET-Re-AnsA (**Kaźmierczyk, praca magisterska**). Genomowy DNA, z którego uzyskano gen *ansA*, został wyizolowany z bakterii *Rhizobium etli* szczep CFN 42 pochodzących z Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ).

Dalekosiężnym celem tego projektu jest poznanie pierwszej struktury krystalograficznej asparaginazy typu *Rhizobium etli*. Moje zadania w tym wieloetapowym przedsięwzięciu obejmowały produkcję rekombinowanego białka, jego oczyszczanie i krystalizację. Literaturowe doniesienia sugerują, że struktura enzymu może znacznie różnić się od struktur zbadanych dotąd asparaginaz typu bakteryjnego i roślinnego ze względu na odmienną sekwencję. Biorąc pod uwagę, że metoda podstawienia cząsteczkowego w tym przypadku może okazać się nieskuteczna, w moich badaniach przyjęłam strategię, która umożliwia wykorzystanie sygnału anomalnego do rozwiązywania struktury krystalicznej.

6.2. Część doświadczalna

6.2.1. Badania wstępne

Przejmując projekt i wektor pET-Re-AnsA, w pierwszej kolejności wykonałam wstępną analizę bioinformatyczną dla monomeru rekombinowanego białka AnsA na podstawie jego sekwencji aminokwasowej korzystając z narzędzia ProtParam serwera ExPASy (**Gasteiger *et al.*, 2005**). Monomer AnsA (ok. 39 kDa) zawiera osiem reszt cysteinowych w sekwencji o długości 367 reszt aminokwasowych. Ich położenie w strukturze przestrzennej nie jest znane. Aby uniknąć niespecyficznej agregacji białka poprzez międzycząsteczkowe mostki S-S, wskazane jest stosowanie czynnika redukującego podczas oczyszczania białka. Monomer AnsA zawiera także trzynaście reszt metioninowych, które jako seleno-metionina mogą zostać wykorzystane do rozwiązania struktury przy użyciu sygnału anomalnego metodą MAD lub SAD.

Rysunek 6.1. Porównanie sekwencji aminokwasowej AnsA z bakteryjnymi przedstawicielami asparaginaz typu bakteryjnego (EcAI i EcAII) i typu roślinnego (EcAIII). Reszty katalityczne EcAI i EcAII (TKD) oraz EcAIII (T) zaznaczyłam kolorem czerwonym. Cztery wysoce zachowawcze motywy sekwencyjne występujące u wszystkich asparaginaz typu *R etli* (Rozdział 3.3.3) zaznaczyłam kolorem niebieskim. Gwiazdki oznaczają reszty identyczne, kropki i dwukropki reszty odpowiednio mniej i bardziej podobne. Porównanie wykonałam w programie *ClustalO* (**McWilliam *et al.*, 2013**).

```
AnsA   MTPSEDFVVTDRGGIVENSHRVHAAVVDAKGRLLYA------LGNPTRMTLA--------  46
EcAI   MQKKSIYV-AYTGGTIGMQ-RSEQGYIPVSGHLQRQLALMPEFHRPEMPDFTIHEYTPLM  58
       *  .. :* :  ** :  . *  . . : ..*:*         : .*    ::
AnsA   -RSAAKPAQALAILE--------------TEGVAGYGFDDADIALMCASHSSEDRHIART  91
EcAI   DSSDMTPEDWQHIAEDIKAHYDDYDGFVILHGTDTMAYTASALSFMLENLG--------K 110
         *  .* :   * *               .*.   .:  : :::*  . .        .
AnsA   RAMLSKIKAEEADLRCGGHPSLSEMVNRSWIKQDFIPTAVCSNCSGKHVGMLAGARAIGA 151
EcAI   PVIVTGSQIPLAELRSDGQIN---LLNALYVAANYPINEVTLFF---NNRLYRGNRTTKA 164
        .:::  :   *:**. *: .   ::*  ::  ::  . *       .  :  * *:  *
AnsA   GTDGY---HLPDHPMQGRVKRTVAELCDLDAGDVEWGTDGCNLPTPA--FPLDRLGRIYA 206
EcAI   HADGFDAFASPNLPPLLEAG---IHIRRLNTPPAPHGE-GELIVHPITPQPIG-VVTIYP 219
        :**:     *: *   ..     .:  *::  .  *  *  :  *    *:  :  **
AnsA   KLASAA------------DGSDAGEGQSTRCAALAHIFRAMARHPEMVAGEGRYCTMLMR 254
EcAI   GISADVVRNFLRQPVKALILRSYGVGNAPQNKAFLQELQ-------EASDRGIVVV---- 268
        ::: .               . * *:: :  *: : ::        .: .*   .
AnsA   AFDGALVGKLGADASYAIGVRASDATRQLGTDGALGISVKIEDGNLEMLYAVVTELLERL 314
EcAI   NLTQCMSGKVNM-GGYATGN----ALAHAGVIGGADMTV--EATLTKLHYLLSQEL-DT- 319
        :  .: **:   ..** *     *  : *. *.  ::*  *    :: * :  ** :
AnsA   GIGSPDVRSQLASFHHPQRVNTMGVTTGGVSFPFKLRGSKSNVDDPRLAAVAR 367
EcAI   ----ETIRK---------------------AMSQNLRGELTPDD--------- 338
            :*.                     ::  :***. :  *


AnsA   MTPSEDFVVTDRGGIVENSHRVHAAVVDAKGRLLYALGNPTRMTLARSAAKPAQALAILE  60
EcAII  ------------------------------------------------------LPNITILA   8
                                                                ::**
AnsA   TEGVAGYGFDDADIALMCASHSSEDRHIARTRAMLS---KIKAEEADLRCGGHPSLSEMV 117
EcAII  TGGTIAGGGDSATKSNYTVGKVGVE-NLVNAVPQLKDIANVKGEQ-VV----NIGSQDMN  62
       * *. . * *.*  :   ..: . : .:..:   *.   ::*.*:  :    . . .:*
AnsA   NRSWIKQDFIPTAVCSNCSGKHVGMLAGARAIGAGTDGYHLPDHPMQGRVKRTVAELCDL 177
EcAII  DNVWLTLAKKINTDCDKTDGFVI---------THGTDTMEETAY------FLDLTVKCDK 107
       :. *:.     .: *.: .*  :          ***  .   :         ::  **
AnsA   DAGDVEWGTDGCNLPTPAFPLDRLGRIYAKLASAADGSDAGEGQSTRC------------ 225
EcAII  PV-----VMVGAMRPSTSMSADGPFNLYNAVVTAADKASANRGVLVVMNDTVLDGRDVTK 162
        .        *.  *: ::  *   .:*  :.:*** :.* .*  .
AnsA   --AALAHIFRAMARHPEMV--AGEGRYCTMLM------RAFDGAL---VGKLGADASYAI 272
EcAII  TNTTDVATFKSVNYGPLGYIHNGKIDYQRTPARKHTSDTPFDVSKLNELPKVGIVYNYAN 222
         :: .  *:::   *      *:  *             ** :    : *:*   .**
AnsA   GVR-ASDATRQLGTDGALGISVKIEDGNLEMLYAVVTELLERLGIGSPDVRSQLASFHHP 331
EcAII  ASDLPAKALVDAGYDGIVSA--GVGNGNL--YKSVFDTLATAAKTGTAVVRSS----RVP 274
       .    :.*  : * ** :.    : :***    :*.  *      *:  ***.    : *
AnsA   QRVNTMGVTTGGVSFPFKLRGSKSNV-------------DDPRL-AAVAR-- 367
EcAII  TGATTQDAEVDDAKYGFVASGTLNPQKARVLLQLALTQTKDPQQIQQIFNQY 326
         ..*  . .  ..: *   *: .               .**:    : .


AnsA   MTPSEDFVVTDRGGIV---ENSHRV---------HAAVVDAKGRLLYALGNPTRMTLARS  48
EcAIII ---MGKAVIAIHGGAGAISRAQMSLQQELRYIEALSAIVET-GQKMLEAGESA-LDV---  52
            . *:: :**     . .  :           :*:*:: *: :   *: : : :
AnsA   AAKPAQALAILETEGVAGYGFDDADIALMCASHSSEDRHIARTRAMLSKIKAEEADLRCG 108
EcAIII ---VTEAVRLLEECPLFNAG-----IGAVFT---RDETHE------LDACVMDGNTLKAG  95
           ::*: :**   :   *     *. : :    :: *       *.    :   *:.*
AnsA   GHPSLSEMVNRSWIKQDFIPTAVCSNCSGKHVGMLAGARAIGAGTDGYHLPDHPMQGRVK 168
EcAIII AVAGVSHLRNP------VLAARLVME-QSPHV------MMIGEGAENFAFARG-----ME 137
       .  .:*.: *       .: : :  : .. **        ** *:: : :        ::
AnsA   RTVAELCDLDA--GD-VEWGTDG-CNLPTPAFPLDRLGRIYA--KLASAADGSDAGEGQS 222
EcAIII RVSPEIFSTSLRYEQLLAARKEGATVLDHSGAPLDEKQKMGTVGAVALDLDGNLAA---A 194
       *.  *: . .    : :   .:*   *   . ***.  :: :   :*   **. *.   :
AnsA   TRCAALAHIFRAMARHPEMVAGEGRYCTMLMRAFDGALVGKLGADASYAIGVRASDATRQ 282
EcAIII TS----------------------------TGGMTNKLPGRVGDSPLVGAGCYANNASVA 226
       *                              .:   * *::* .   . *  *.:*:
AnsA   LGTDG----------ALGISVKIEDGNLEMLYAV---VTELLERLGIGSPDVRSQLASFH 329
EcAIII VSCTGTGEVFIRALAAYDIAALMDYGGLSLAEACERVVMEKLPALGGSG-----GLIAI- 280
       :.  *          *  *:. :: * *.:  *    * * *  ** ..       * ::
AnsA   HPQRVNTMGVTTGGVSFPFKLRGSKSNV----DDPRLAAVAR--------- 367
EcAIII ---------DHEGNVALPFNTEGMYRAWGYAGDTPTT-GIYREKGDTVATQ 321
                  * *::**: .*         * *   .: *
```

Wykonałam także analizę porównawczą sekwencji asparaginazy z *Rhizobium etli* z sekwencjami przedstawicieli trzech innych typów asparaginaz w programie *ClustalO* (**McWilliam *et al.*, 2013**) (Rysunek 6.1). Analiza ta wykazała, dla wszystkich trzech porównań (AnsA/EcAI, AnsA/EcAII i AnsA/EcAIII), że zarówno identyczność jak i podobieństwo sekwencyjne są bardzo niskie i wynoszą odpowiednio ok. 16% i ok. 41%. Potwierdza to literaturowe doniesienia o znacznej odmienności białka AnsA (**Borek i Jaskólski, 2001; Moreno-Enriquez *et al.*, 2012; Huerta-Saquero *et al.*, 2013).**

Wektor ekspresyjny pET-Re-AnsA został przygotowany w oparciu o plazmidowy wektor pET151/D-TOPO firmy INVITROGEN. Konstrukcja wektora daje możliwość prowadzenia indukowanej za pomocą IPTG ekspresji białka na pożywkach selektywnych zawierających ampicylinę. Białko produkowane jest z N-terminalnym znacznikiem histydynowym, którego masę (ok. 4 kDa) należy uwzględnić weryfikując wyniki kolejnych etapów oczyszczania na żelach SDS-PAGE. Istnieje możliwość usunięcia znacznika histydynowego przy użyciu proteazy TEV, przy czym po cięciu na aminowym końcu białka pozostaje sześć reszt aminokwasowych (GIDPFT-). Stąd, masa monomeru AnsA na żelu SDS-PAGE powinna wynosić 43 kDa dla białka ze znacznikiem histydynowym oraz 39.5 kDa dla białka bez znacznika.

Tabela 6.1. Protokół powielenia plazmidowego wektora ekspresyjnego pET-Re-AnsA.

etap	stosowane materiały	procedura
transformacja	wektor pET-Re-AnsA komórki top10 (INVITROGEN) pożywka płynna SOC	- delikatne zmieszanie rozmrożonych komórek z wektorem (ok. 5 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5 min. - zmieszanie komórek z SOC (250 μL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla na pożywce selektywnej	pożywka płynna LB z ampicylina (0.1 mg/mL)	- rozcieńczenie hodowli pożywką LB (3 mL) - inkubacja 37ºC, 16 godz., 250 rpm
izolacja plazmidowych wektorów ekspresyjnych	wykonana zgodnie z protokołem literaturowym (**Sambrook i Russel, 2001**; Protokół nr 1.32)	
weryfikacja wyniku	- pomiar stężenia DNA przy użyciu spektrofotometru NanoDrop	
zabezpieczenie wektorów	- zamrożenie w -20ºC	

Przejmując wektor pET-Re-AnsA zadbałam również o jego zabezpieczenie (Tabela 6.1), ponieważ preparat taki zazwyczaj służy wielokrotnie do różnych eksperymentów, ulegając zużyciu. W tym celu powieliłam wektor ekspresyjny poprzez namnożenie w komórkach *E. coli*, a następnie wyizolowałam go zgodnie z protokołem

literaturowym (**Sambrook i Russel, 2001;** Protokół nr 1.32) uzyskując ok. 50 µg DNA.

Potwierdziłam również, że sekwencja insertu w pET-Re-AnsA jest identyczna z sekwencją genu asparaginazy II z *Rhizobium etli* CFN42 zdeponowaną w bazie NCBI pod numerami indentyfikacyjnymi: Gene ID 3895547 i Protein ID YP_472512.1 (**Gonzalez *et al.*, 2006**). Sekwencjonowanie wykonała firma GENOMED z Warszawy.

6.2.2. Produkcja białka natywnego

Aby określić warunki pozwalające na produkcję rozpuszczalnego białka AnsA, przeprowadziłam kilka pilotażowych eksperymentów. Wykonałam transformację komórek BL21 Star (DE3) firmy INVITROGEN, które wyhodowałam na selektywnej pożywce LB na szalkach Petriego, a następnie wykorzystałam do przygotowania kilku różnych płynnych hodowli komórkowych.

Tabela.6.2. Protokół produkcji natywnego białka AnsA.

etap	stosowane materiały	procedura
transformacja	wektor pET-Re-AnsA komórki BL21 Star (INVITROGEN) pożywka płynna SOC	- delikatne zmieszanie komórek z wektorem (ok. 5 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5 min. - zmieszanie komórek z SOC (250 µL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla na pożywce stałej	pożywka stała LB z ampicyliną (0.1 mg/mL)	- posiew na szalkę z pożywką stałą LB - inkubacja 37ºC, 16 godz.
hodowla na pożywce płynnej	pożywka płynna TB z ampicyliną (0.1 mg/mL)	- przeniesienie komórek z szalki do pożywki płynnej TB (250 mL) - inkubacja 37ºC, 2 godz., 250 rpm - rozcieńczenie hodowli świeżą pożywką TB (do 2 L) - inkubacja 37ºC, 250 rpm, do OD_{600}=0.8 tj. ok. 3 godz. - obniżenie temperatury do 25ºC
ekspresja	IPTG (do stężenia 0.2 mM)	- dodanie IPTG - inkubacja 25ºC, 16 godz., 250 rpm
zabezpieczenie komórek		- wirowanie 5000 rmp, 10ºC, 30 min. - zamrożenie komórek w -80ºC

Przetestowałam dwie temperatury ekspresji (37ºC i 25ºC), dwa rodzaje pożywek (LB i TB) oraz dwa stężenia induktora ekspresji IPTG (0.2 mM i 0.4 mM). Sprawdzałam również, czy czas prowadzenia ekspresji wpływa na obecność białka we frakcjach rozpuszczalnych i nierozpuszczalnych. Eksperymenty prowadzone w temperaturze 37ºC pozwalały uzyskać mniej więcej równe ilości białka w obu

frakcjach, natomiast obniżenie temperatury ekspresji do 25ºC spowodowało zmianę proporcji białka AnsA we frakcji rozpuszczalnej i nierozpuszczalnej na korzyść tej pierwszej. Podobnie niższe stężenie induktora powodowało wolniejszą produkcję rekombinowanego białka i zwiększało pulę białek rozpuszczalnych. Wpływ pożywki był znikomy.

Opierając się na doświadczeniu zdobytym podczas pracy z białkami Nod, przygotowałam protokół produkcji natywnego białka AnsA na dużą skalę (Tabela 6.2). Białko produkowałam prowadząc ekspresję w temperaturze 25ºC przez ok. 16 godzin, indukując ją za pomocą IPTG o stężeniu 0.2 mM. Wydajność ekspresji wnosi ok. 16 g osadu komórkowego z 1 L hodowli.

6.2.3. Produkcja białka znakowanego atomami selenu

Do produkcji białka AnsA znakowanego atomami selenu (AnsA-Se) wykorzystałam metodę polegającą na blokowaniu szlaku biosyntezy metioniny w komórkach *E. coli*, oraz pożywkę minimalną. Tym razem nie robiłam wstępnych eksperymentów na małą skalę, starając się od razu wyprodukować białko w ilościach umożliwiających jego oczyszczanie.

Początkowo opierałam się na protokole opracowanym przez dr Krzysztofa Brzezińskiego (**Brzeziński, 2006**). Postępowałam ściśle według przepisu, przy czym do transformacji wykorzystałam komórki BL21 Star (DE3) firmy INVITROGEN, w których z powodzeniem prowadziłam produkcję natywnego białka AnsA. W skrócie postępowanie było następujące. Hodowlę komórek prowadziłam w temperaturze 30ºC, początkowo w pożywce LB, później w pożywce minimalnej (MDL), stopniowo zwiększając objętość hodowli poprzez dodawanie świeżych pożywek z ampicyliną. W trakcie dwudniowej hodowli, przez jedną noc bakterie pozostawały na lodzie. Gdy wartość OD_{600} osiągnęła 0.8 temperaturę hodowli obniżyłam, dodałam kolejno roztwory MetStop (**Van Duyne *et al.*, 1993**) oraz Se-Met (do stężenia 50 µg/mL) (MDL) i odczekałam 30 minut. Dodałam IPTG indukując ekspresję, którą prowadziłam w warunkach ustalonych dla białka natywnego, czyli przez 16 godzin w 25ºC. Z 1 L hodowli otrzymałam zaledwie 4 g osadu komórkowego. Niestety nie udało mi się uzyskać białka AnsA-Se z tej ekspresji z powodu strat w trakcie oczyszczania.

Tabela 6.3. Ostateczny protokół produkcji białka AnsA znakowanego selenometioniną (AnsA-Se).

etap	stosowane materiały	procedura
transformacja	wektor pET-Re-AnsA komórki BL21 Star (DE3) (INVITROGEN) pożywka płynna SOC	- delikatne zmieszanie rozmrożonych komórek z wektorem (ok. 5 ng) - inkubacja na lodzie, 30 min. - inkubacja 42ºC, 30 sek. - inkubacja na lodzie, 5 min. - zmieszanie komórek z SOC (250 µL) - inkubacja 37ºC, 1 godz., 200 rpm
hodowla na pożywce stałej	pożywka stała LB z ampicyliną (0.1 mg/mL)	- posiew na szalkę z pożywką stałą LB - inkubacja 37ºC, 16 godz.
hodowla na pożywce płynnej	pożywka płynna TB z ampicyliną (0.1 mg/mL) pożywka minimalna (MDL) z ampicyliną (0.1 mg/mL)	- przeniesienie komórek z szalki do pożywki płynnej TB (250 mL) - inkubacja 37ºC, 4 - 6 godz., 250 rpm - zwirowanie komórek (5000 rmp, 10ºC, 10 min.) - odrzucenie pożywki - rozpuszczenie osadu komórkowego w świeżej pożywce minimalnej (2 L) - inkubacja 30ºC, 250 rpm, do OD_{600}=0.8 czyli ok. 8 - 10 godz. - obniżenie temperatury do 18ºC
zahamowanie biosyntezy metioniny	roztwór MetStop: 1% (w/v) L-lizyna 1% (w/v) L-treonina 1% (w/v) L-fenyloalanina 0.5% (w/v) L-walina 0.5% (w/v) L-izoleucyna 0.5% (w/v) L-leucyna roztwór Se-Met (10 mg/mL) (MDL)	- dodanie MetStop (20 mL / 2 L pożywki) - dodanie Se-Met (100 mg / 2 L pożywki) - inkubacja 18ºC, 250 rpm, 30 min.
ekspresja	IPTG (do stężenia 0.1 mM)	- dodanie IPTG - inkubacja 25ºC, 18 godz., 250 rpm
zabezpieczenie komórek		- wirowanie 5000 rmp, 10ºC, 30 min. - zamrożenie komórek w -80ºC

Z powodu niskiej wydajności zastosowanego protokołu, w drugim podejściu dokonałam jego modyfikacji (Tabela 6.3). Przede wszystkim prowadziłam hodowlę bez przerwy, pomijając etap w którym komórki pozostawały przez noc na lodzie, aby uniknąć ich obumierania. Zmieniłam pożywkę LB na TB, którą stosowałam podczas produkcji białka natywnego. Hodując komórki w pożywce TB stosowałam temperaturę 37ºC aby przyspieszyć ich namnażanie. Zachowałam temperaturę 30ºC podczas hodowli na pożywce minimalnej. Wprowadziłam etap krótkiego i delikatnego wirowania aby uniknąć mieszania się pożywek TB i minimalnej. Etap zahamowania biosyntezy metioniny pozostał bez zmian. Ekspresję prowadziłam w temperaturze 25ºC, tj. tak jak ekspresję białka natywnego, przy czym zmniejszyłam ilość dodawanego induktora ekspresji. Zmieniony protokół wymagał 14 - 18 godzin pracy podczas hodowli, od momentu pobierania kolonii z szalki Petriego do uruchomienia ekspresji.

Cała procedura pozwoliła przyspieszyć proces produkcji białka, w którym hodowla i ekspresja wymagała łącznie ok. dwóch dni. Z 1 L hodowli przy użyciu zmodyfikowanego protokołu otrzymywałam ok. 12 g osadu komórkowego, co stanowi wynik zbliżony do wyniku ekspresji białka natywnego.

6.2.4. Oczyszczanie białka AnsA

Przystępując do oczyszczania natywnego białka AnsA, opracowałam protokół składający się z następujących etapów: liza, chromatografia powinowactwa, cięcie proteazą TEV^{His} i chromatografia powinowactwa. Zależało mi, aby protokół pozwalał na szybkie (maksymalnie dwudniowe) oczyszczanie. Postanowiłam także używać roztworów o małej ilości składników, które w razie potrzeby mogłabym łatwo modyfikować. Podstawowe składniki roztworów używanych do oczyszczania białka AnsA to: bufor (20 mM Tris pH 7.5), sól (200 mM NaCl), czynnik redukujący (5 mM β-merkaptoetanol) oraz czynnik wymywający ze złoża niklowego (20 - 300 mM imidazol).

Etap lizy składał się z trzech kroków: trawienia lizozymem (30 min.), lizy ultradźwiękami (sonikacji), dalszego trawienia lizozymem (30 min.). Później etap ten zmodyfikowałam pozwalając na dłuższe trawienie lizozymem, co umożliwiało izolację większej ilości rozpuszczalnego białka. Frakcje białek rozpuszczalnych i nierozpuszczalnych rozdzielałam przez wirowanie.

Użycie chromatografii powinowactwa podyktowała konstrukcja wektora ekspresyjnego, który pozwalał produkować białko w połączeniu ze znacznikiem histydynowym. Wykorzystywałam w związku z tym złoże Ni Sepharose High Performance (GE Healthcare) ładowane $NiSO_4$ do oczyszczania na kolumnie metodą grawitacyjną. Szczególnie ważny był moment wiązania białka AnsA do złoża. Im dłużej trwał, tym lepsze wydajności uzyskiwałam. Z kolei 20 mM imidazol obecny w roztworze rozpuszczalnych białek pozwalał uniknąć ich niespecyficznego wiązania się do kolumny. Jedną z zalet złoża Ni Sepharose High Performance (GE Healthcare) było to, że pozwalało ono na użycie podciśnienia, co znacznie przyspieszało procedurę oczyszczania. Podciśnienie stosowałam jednak tylko na etapie intensywnego, wielokrotnego przemywania związanych białek. Elucję prowadziłam przy użyciu gradientu 50 - 300 mM imidazolu.

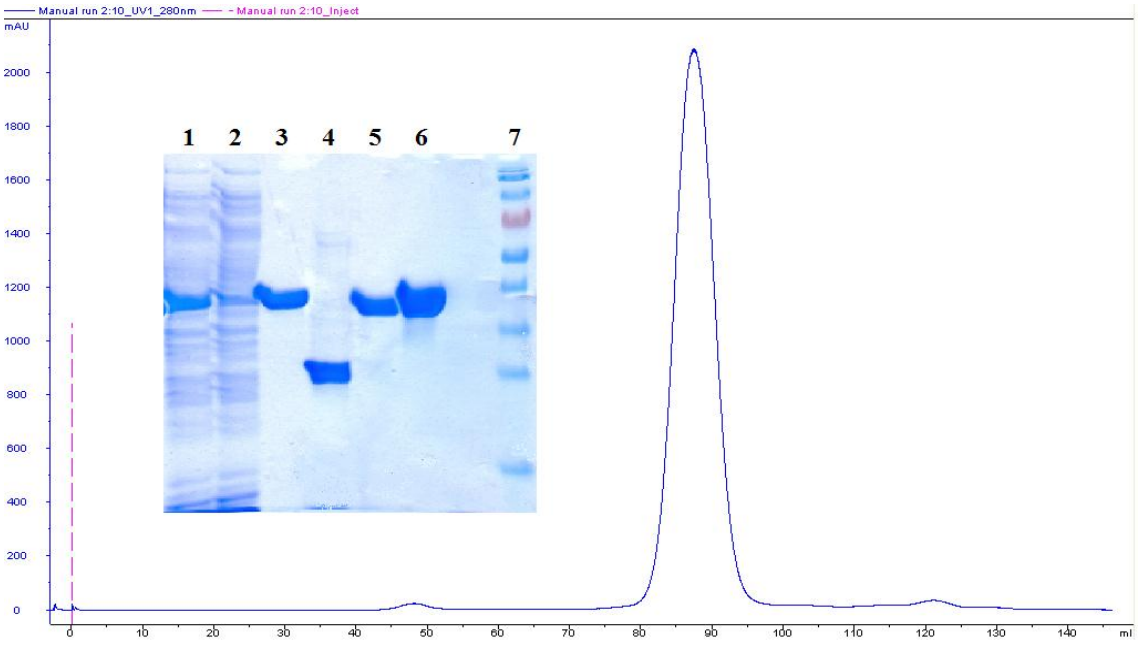
Cięcie proteazą TEV^{His} (rekombinowanym enzymem zawierającym również

znacznik histydynowy) prowadziłam podczas dializy, którą stosowałam aby usunąć imidazol używany do wymywania białka ze złoża niklowego. Połączenie tych dwóch etapów znacząco skracało czas oczyszczania. Etap ten prowadziłam przez noc.

Drugą chromatografię powinowactwa prowadziłam używając tego samego złoża niklowego. Miała ona na celu oddzielenie białka AnsA pozbawionego znacznika histydynowego od wiązanych do złoża białek AnsA nadal posiadających znacznik histydynowy, odciętych znaczników histydynowych i proteazy TEV^{His}, którą dodawałam w poprzednim etapie pracy. Dopracowując protokół ponownie wydłużyłam etap wiązania do złoża, aby zwiększyć wydajność.

Pierwsze oczyszczanie prowadziłam powoli, sprawdzając w której frakcji wypływa białko AnsA i oceniając jego czystość za pomocą elektroforezy SDS-PAGE. Początkowo czystość białka nie była zadowalająca, więc protokół wydłużyłam, dodając na końcu etap chromatografii żelowej. Rozdział prowadziłam przy użyciu FPLC i systemu AKTA na kolumnie GE HiLoad 16/60 Superdex 200. Białko AnsA wypłynęło z kolumny ok. 87 mL po iniekcji (Rysunek 6.2) co może oznaczać, że w roztworze jest dimerem. Oszacowałam to na podstawie krzywej kalibracyjnej przygotowanej w oparciu o różne białka o znanej masie.

Rysunek 6.2. Wynik z poszczególnych etapów oczyszczania białka AnsA, pokazany na chromatogramie FPLC i żelu SDS-PAGE: 1- etap lizy, supernatant uzyskany po wirowaniu; 2 i 3 – etap chromatografii powinowactwa, 2- frakcja białek nie związanych do złoża, 3 – eluat; 4 – proteaza TEV^{His}, 5 – etap drugiej chromatografii powinowactwa, białko AnsA bez znacznika histydynowego; 6 – etap chromatografii żelowej. Pik na chromatogramie to białko AnsA wypływające z kolumny ok. 87 mL od iniekcji.



Już drugie oczyszczanie pozwoliło uzyskać czysty preparat białka AnsA w ilości ok. 10 mg z 4 g osadu komórkowego (ok. 250 mL hodowli) nadające się do krystalizacji. Późniejsze modyfikacje, polegające jedynie na dopracowywaniu objętości używanych roztworów i czasu trwania poszczególnych kroków, pozwoliły wypracować protokół, dzięki któremu otrzymywałam nawet 14 mg czystego białka AnsA z 4 g osadu komórkowego (ok. 250 mL hodowli) (Tabela 6.4).

Tabela 6.4. Protokół oczyszczania rekombinowanego białka AnsA

etap	stosowane odczynniki	procedura
liza	roztwór 1: 50 mM Tris pH 7.5 200 mM NaCl 5 mM β-merkaptoetanol 20 mM imidazol PMSF (do stężenia 1 mM) lizozym (do stężenia 1 mg/mL)	- rozpuszczenie osadu komórkowego w roztworze 1 (ok. 30 mL / 4 g osadu) - dodanie PMSF i lizozymu - 1godz. delikatnego mieszania w 20ºC - sonikacja (18 x 10 sekund, moc 80%) - 1godz. delikatnego mieszania w 20ºC - wirowanie: 20000 rpm, 10ºC, 30 min. - odrzucenie osadu
chromatografia powinowactwa na kolumnie niklowej metodą grawitacyjną	złoże Ni Sepharose High Performance (GE-Healthcare) załadowane $NiSO_4$ i przemyte roztworem 1 roztwór 2: 50 mM Tris pH 7.5 200 mM NaCl 5 mM β-merkaptoetanol 300 mM imidazol	- zmieszanie supernatantu ze złożem wypełniającym kolumnę (ok. 10 mL złoża / 30 mL supernatantu) - inkubacja 30 min, temp. 20ºC - odrzucenie eluatu zawierającego frakcję białek niezwiązanych do złoża - przemywanie roztworem 1 (10 x 30 mL) z użyciem podciśnienia - wymywanie białka roztworem 2 (ok. 20 mL)
usunięcie znacznika histydynowego	TEV^{His} (0.15 mg / 1 mg białka)	- zmieszanie eluatu z proteazą TEV^{His} - dializa białek do roztworu 1 (16 godz., 4ºC, delikatne mieszanie)
chromatografia powinowactwa na kolumnie niklowej metodą grawitacyjną	złoże Ni Sepharose High Performance (GE-Healthcare) załadowane $NiSO_4$ i przemyte roztworem 1	- zmieszanie roztworu białek ze złożem wypełniającym kolumnę (ok. 10 mL złoża / 20 mL roztworu) - inkubacja 20 min., temp. 20ºC - zebranie eluatu zawierającego białko AnsA
chromatografia żelowa przy użyciu FPLC	kolumna GE HiLoad 16/60 Superdex 200 (GE-Healthcare) roztwór 3: 50 mM Tris pH 7.5 200 mM NaCl 5 mM β-merkaptoetanol	- zatężenie eluatu przy użyciu filtrów wirówkowych Amicon (MILLIPORE) - nałożenie preparatu na kolumnę - rozdział - połączenie frakcji zawierających AnsA
zmiana składu roztworu białka	roztwór 4: 20 mM Tris pH 7.5 3 mM TCEP	- zatężenie preparatu przy użyciu filtrów wirówkowych Amicon (MILLIPORE) - dializa do roztworu 4

Ostatni etap dopracowanego protokołu (Tabela 6.4), tj. zmiana składu roztworu białka, został dodany podczas poszukiwań wstępnych warunków krystalizacji i jest opisany w kolejnym rozdziale. Oczyszczone i zatężone białko poddawałam testom przy użyciu DLS, aby sprawdzić jego jakość (Rozdział 6.2.6), a następnie świeży preparat służył wprost do krystalizacji lub był przechowywany w 4 i -80ºC.

Dopracowany protokół (Tabela 6.4) sprawdzał się także podczas oczyszczania białka znakowanego selenometioniną (AnsA-Se). W tym przypadku uzyskiwałam maksymalnie ok. 10 mg białka AnsA-Se z 4 g osadu komórkowego.

6.2.5. Krystalizacja białka AnsA

Podczas pracy z asparaginazą z *Rhizobium etli* najwięcej czasu i energii poświęciłam na próby krystalizacyjne. Składały się na nie, na przemian, poszukiwania wstępnych warunków, optymalizacje krystalizacji oraz testy rozpraszania rentgenowskiego, wykonywane przy użyciu promieniowania synchrotronowego. Rozpoczynając prace nie posiadałam żadnych wskazówek co do potencjalnych warunków krystalizacji białka AnsA.

6.2.5.1. Poszukiwanie wstępnych warunków krystalizacji

Podczas poszukiwania wstępnych warunków krystalizacji korzystałam z robota krystalizacyjnego i komercyjnych zestawów do krystalizacji firm Hampton Research i MDL. Użycie robota pozwalało przetestować dużą liczbę warunków, przy jednoczesnej oszczędności odczynników, preparatu białka i czasu. Próby były prowadzone w temperaturze 19ºC.

Rozpoczynając prace krystalizacyjne posiadałam natywne białko AnsA oczyszczone według wstępnego protokołu, który nie zawierał ostatniego etapu, czyli wymiany buforu. W związku z tym pierwsze partie preparatu znajdowały się w buforze używanym do filtracji żelowej (Tabela 6.5 preparat A) i były zatężone do ok. 15 mg/mL. Sumarycznie preparaty te wystarczyły do nastawienia krystalizacji z ośmioma komercyjnymi zestawami roztworów, czyli łącznie dla 768 warunków. Dla każdego warunku nastawiałam trzy krople w proporcjach: 0.1:0.2, 0.2:0.2 i 0.2:0.1 µL białka do rezerwuaru. Pierwsze próby krystalizacyjne nie przyniosły pozytywnych wyników, krople pozostawały klarowne lub zawierały strącone białko.

W kolejnej rundzie krystalizacji wykorzystywałam preparat o stężeniu ok. 15 mg/mL, który po oczyszczaniu był dializowany w celu całkowitego usunięcia soli (Tabela 6.5 preparat B). Tym razem nastawiłam po jednej kropli (w proporcji 0.3:0.3 μL) wykorzystując 11 komercyjnych zestawów roztworów, czyli łącznie 1056 warunków. Usunięcie NaCl okazało się zabiegiem kluczowym, ponieważ już po kilku dniach w kroplach pojawiły się pierwsze drobne kryształy w dwóch warunkach zawierających fosforany. Wyselekcjonowane warunki optymalizowałam, jednak wyniki nie były zadowalające.

Chcąc znaleźć nowe warunki krystalizacji, z których kryształy nadawałyby się do pomiarów dyfrakcji rentgenowskiej, wróciłam do przeszukiwania standardowych warunków w komercyjnych zestawach do krystalizacji i za pomocą robota krystalizacyjnego nastawiłam łącznie ok. 3.5 tys. nowych kropel. Ponownie zmieniałam skład preparatu białka poprzez dializę wykonywaną jako ostatni etap oczyszczania. Przygotowałam w ten sposób całą serię nowych preparatów do przetestowania (Tabela 6.5 preparaty D, E, F, G, H, I, J, K). Zatężałam preparaty do ok 13 mg/mL, nastawiałam krystalizację i umieszczałam płytki w 19ºC. Efekty procesu krystalizacji sprawdzałam regularnie po 1 dniu, po 7 dniach, po 2 tygodniach i po miesiącu od nastawienia. Tak jak poprzednio, kryształy rosły w warunkach zawierających fosforany, ale pojawiły się także w nowych warunkach zawierających octany i siarczany.

Tabela 6.5. Składniki poszczególnych preparatów białka, które poddawałam krystalizacji za pomocą robota i komercyjnych roztworów do krystalizacji („10x" oznacza 10-krotny nadmiar molowy).

A 15 mg/mL białko AnsA 20 mM Tris pH 7.5 200 mM NaCl 5 mM β-merkaptoetanol	**B** 15 mg/mL białko AnsA 20 mM Tris pH 7.5 5 mM β-merkaptoetanol	**C** 16 mg/mL białko AnsA 20 mM $(K_2H/NaH_2)PO_4$ pH 9.0 5 mM TCEP	
D 13 mg/mL białko AnsA 20 mM Tris pH 7.5 5 mM DDT	**E** 13 mg/mL białko AnsA 20 mM Tris pH 7.5 5 mM TCEP	**F** 13 mg/mL białko AnsA 20 mM Tris pH 7.5	**G** 13 mg/mL białko AnsA 20 mM Tris pH 7.5 5 mM TCEP 5% glicerol
H 13 mg/mL białko AnsA 20 mM Tris pH 7.5 5 mM DDT (10x) K-Asp	**I** 13 mg/mL białko AnsA 20 mM Tris pH 7.5 5 mM TCEP (10x) K-Asp	**J** 13 mg/mL białko AnsA 20 mM Tris pH 7.5 (10x) K-Asp	**K** 13 mg/mL białko AnsA 20 mM Tris pH 7.5 5 mM TCEP 5% glicerol (10x) K-Asp

Podsumowując efekty poszukiwania wstępnych warunków krystalizacji za pomocą serii preparatów białka mogę stwierdzić, że zastosowane odczynniki dawały

różne efekty. Dodanie glicerolu zwiększyło rozpuszczalność białka i zmniejszyło częstotliwość pojawiania się kryształów (Tabela 6.5 preparaty G i K). Dodanie do preparatu białka soli kwasu asparaginowego, czyli produktu katalizowanej reakcji, nie miało znaczącego wpływu na krystalizację (Tabela 6.5 roztwory H, I, J, K). Usunięcie czynnika redukującego (Tabela 6.5 preparat F i J) powodowało, że białko stawało się mniej stabilne, wytrącało się podczas zatężania i znacznie częściej w trakcie krystalizacji. Rodzaj czynnika redukującego generalnie nie zmieniał ani pokroju kryształów, ani warunków, w których występowały. Po pewnym czasie zauważyłam jednak, że krople zawierające TCEP (Tabela 6.5 roztwory E, I, G, K) były dłużej stabilne. Dlatego do ponownej optymalizacji warunków fosforanowych oraz optymalizacji nowo znalezionych warunków siarczanowych i octanowych wybrałam preparaty E i I (Tabela 6.5). Warunki fosforanowe, siarczanowe i octanowe oraz ich optymalizację opisuję w kolejnym rozdziale.

Charakterystyka biochemiczna AnsA (**Moreno-Enriquez *et al.*, 2012**) pokazała, że białko jest najaktywniejsze w pH 9.0. Zatem do poszukiwania wstępnych warunków krystalizacji wykorzystałam także preparat w buforze, w którym wykonano wspomniany test, tj. w buforze fosforanowym (Tabela 6.5 preparat C). Miałam nadzieje, że zmiana pH preparatu wpłynie na jakość dyfrakcji rentgenowskiej kryształów. W tym przypadku przetestowałam jedenaście zestawów roztworów do krystalizacji (1056 warunków) w trzech proporcjach białka do rezerwuaru: 0.2:0.4, 0.3:0.3 i 0.4:0.2 μL. Tym razem kryształy białka nie pojawiły się w warunkach fosforanowych, natomiast warunki octanowe i siarczanowe mogłam ponownie optymalizować.

6.2.5.2. Optymalizacja krystalizacji

Kryształy białka pochodzące z malutkich kropel nastawianych przez robota często były bardzo drobne i rosły w dużym zagęszczeniu, co uniemożliwiało ich wyłowienie. Inne były bezkształtne lub pozrastane i, co najważniejsze, nie rozpraszały promieniowania rentgenowskiego. Dlatego konieczna była optymalizacja warunków krystalizacji.

W pierwszej kolejności nastawiałam tzw. powtórki warunków wstępnych. Były to ręcznie przygotowywane krople zawierające oryginalny roztwór do krystalizacji z zestawu komercyjnego oraz ten sam preparat białka co stosowany podczas krystalizacji przesiewowych. Krople te miały znacznie większą objętości niż krople nastawiane przez robota i były kroplami wiszącymi, stąd ten układ krystalizacyjny nie był w pełni

identyczny z sytuacją oryginalną. Krople nastawiałam w trzech proporcjach preparatu białka do roztworu krystalizacyjnego. W proporcji identycznej jak w eksperymencie przesiewowym (np. 1:1), oraz w proporcjach nieznacznie zmienionych (np. 1:2 i 2:1).

Większość warunków jakie znalazłam podczas zakrojonych na szeroką skalę krystalizacji przesiewowych udało mi się powtórzyć. W sumie optymalizowałam wszystkie trzy typy warunków (warunki fosforanowe, siarczanowe i octanowe), jednak ograniczając się do czterech rodzajów preparatu białka: B, E, I oraz C.

Optymalizacja warunków fosforanowych

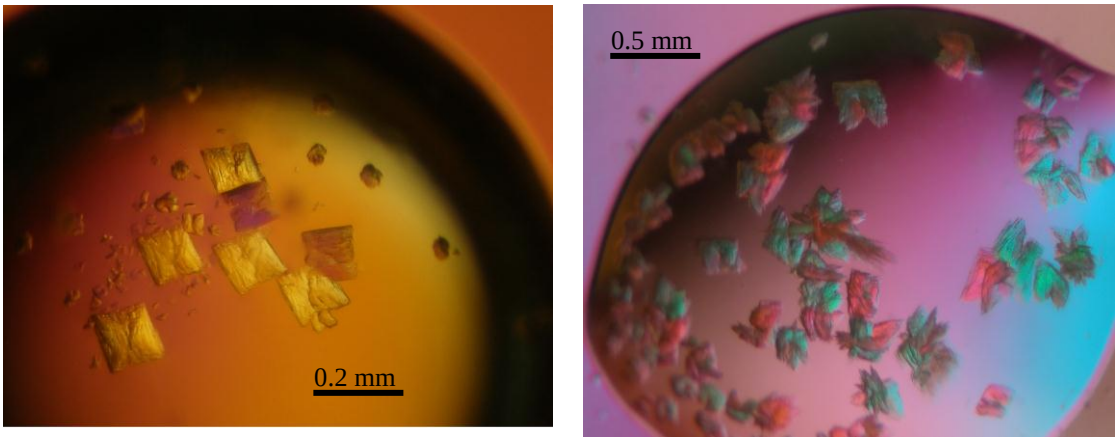
Prawie wszystkie preparaty białka AnsA krystalizowały w dwóch warunkach zawierających fosforany (Tabela 6.6). Optymalizację tych warunków prowadziłam jednak dla preparatów B, E i I, za każdym razem postępując w podobny sposób.

Kryształy z powtórek warunków fosforanowych (Rysunek 6.3) były pokroju płytkowego, nieregularnie pozrastane i bardzo słabo rozpraszały promieniowanie rentgenowskie. Pomiar danych dyfrakcyjnych nie był możliwy, udało się jedynie potwierdzić, że były to kryształy białka.

Tabela 6.6. Warunki fosforanowe krystalizacji białka AnsA.

nazwa zestawu do krystalizacji (kod) producent numer warunku	skład roztworu
INDEX HT (HR2-134) Hampton Research 18 (B6)	0.49 M NaH_2PO_4 0.91 M K_2HPO_4 pH 6.9
INDEX HT (HR2-134) Hampton Research 19 (B7)	0.056 M NaH_2PO_4 1.344 M K_2HPO_4 pH 8.2

Rysunek 6.3. Przykłady kryształów białka AnsA z warunków fosforanowych.



Rozpoczynając optymalizację warunków fosforanowych przygotowałam dwa stężone (4 M) roztwory: NaH_2PO_4 i K_2HPO_4, które, łączone w różnych proporcjach, służyły mi do przygotowywania mieszanin buforowych o różnym pH. Ponadto manipulowałam stężeniami fosforanów w roztworach rezerwuarowych, stężeniami preparatu białka, proporcjami składników kropli krystalizacyjnej i temperaturą krystalizacji (Tabela 6.7). Rozmaite roztwory, każdy w kilku stężeniach, służyły mi jako dodatki do warunków fosforanowych. Stosowałam dwa rodzaje dodatków. Jedne były roztworami zwiększającymi rozpuszczalność, a więc i stabilność białka, co miało spowolnić krystalizację. Drugi rodzaj dodatków (PEG) wybrałam na podstawie obserwacji niefosforanowych warunków, w których krystalizowało białko AnsA. Próbowałam także szczepić krople zarodkami krystalizacyjnymi używając różnych technik wykonania tego zabiegu. Za każdym razem podczas przygotowywania płytki krystalizacyjnej modyfikowałam dwa lub więcej z wymienionych elementów. W sumie, biorąc pod uwagę wszystkie preparaty białka AnsA, nastawiłam ręcznie ok. 700 kropel krystalizacyjnych bazujących na warunkach fosforanowych.

Tabela 6.7. Optymalizacja warunków fosforanowych. Wymieniam zmienne elementy i te ich wartości, które były stosowane, jednostki podaję w nawiasach.

pH roztworów rezerwuarowych		9.0 8.4 8.2 7.6 7.4 7.2 6.9 6.5 6.4 6.2 6.0
stężenia fosforanów [M]		0.3 0.6 0.7 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2
stężenia preparatu białka [mg/mL]		5 7 10 13 15 20 25
proporcje składników kropli [µL]		1:5 2:3 1:3 2:2 3:3 3:1 3:2 5:1
temperatura [ºC]		4 19 30
dodatki do roztworów rezerwuarowych	glicerol [%]	3 6 12
	glicyna [M]	0.1 0.2 0.4
	oktyloglukoza [%]	0.5 1.0 2.0
	dioksan [%]	3 6 12
	PEG 400 [%]	5 10 20
	PEG 3350 [%]	15 20 25 35
	PEG 2K MME [%]	15 20 25 35
	PEG 550 MME [%]	25 35 40 45
	MPD [%]	35 40 45 55

Poza ręcznie nastawianymi kroplami, optymalizacje prowadziłam także przy użyciu robota krystalizacyjnego. Za jego pomocą nastawiłam dwie płytki krystalizacyjne, różniące się zawartością roztworów fosforanowych (Tabela 6.6), ale zawierające te same odczynniki dodatkowe pochodzące z komercyjnego zestawu do krystalizacji Additive Screen HT (96 warunków) firmy Hampton Research.

Niestety, wszystkie kryształy z warunków fosforanowych, niezależnie od użytych preparatów białka i zastosowanych optymalizacji, zawsze rosły jako pozrastane bardzo cienkie płytki i nie nadawały się do pomiarów dyfrakcyjnych.

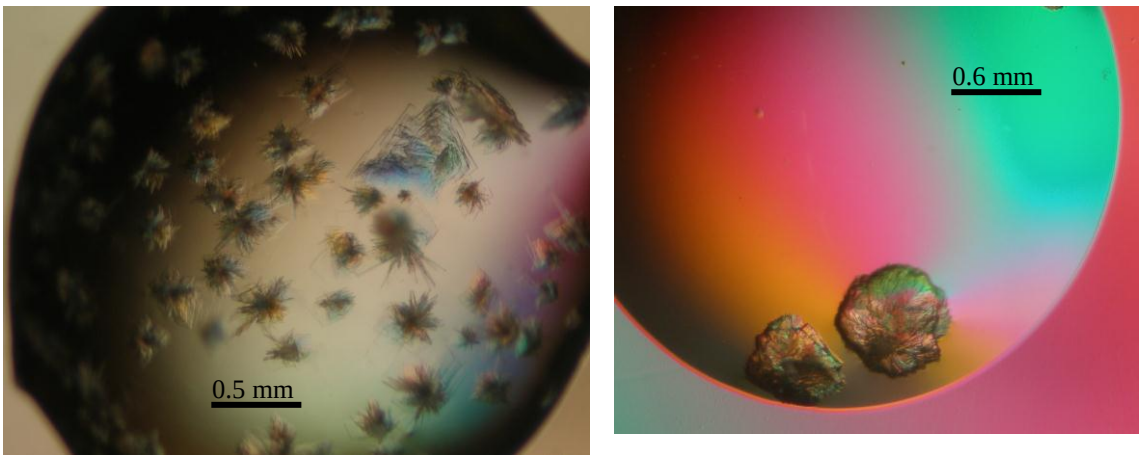
Optymalizacja warunków siarczanowych

Optymalizację warunków siarczanowych (Tabela 6.8) prowadziłam dla preparatów białka C, E i I. Kryształy z powtórek warunków siarczanowych były jeszcze bardziej nieregularne od kryształów z warunków fosforanowych. Przypominały zrosty igieł (w przypadku zastosowania $(NH_4)_2SO_4$), lub też igieł i płytek (w przypadku zastosowania Li_2SO_4) (Rysunek 6.4). Odłamane fragmenty płytek bardzo słabo rozpraszały promieniowanie rentgenowskie, potwierdzając jedynie, że ich budulcem jest białko, natomiast zbudowane z igieł sferyczne zrosty rozpadały się podczas próby wyłowienia, uniemożliwiając pomiar.

Tabela 6.8. Warunki siarczanowe krystalizacji białka AnsA.

nazwa zestawu do krystalizacji (kod) producent numer warunku	skład roztworu
Structure Screen 1 (MD1-01) MDL 32	2.0 M $(NH_4)_2SO_4$ 0.1 M Tris pH 8.5
Structure Screen 1 (MD1-01) MDL 35	0.2 M Li_2SO_4 0.1 M Tris pH 8.5 30% w/v PEG 4K
INDEX HT (HR2-134) Hampton Research 77 (G5)	0.2 M Li_2SO_4 0.1 M Tris pH 8.5 25% w/v PEG 3350

Rysunek 6.4. Przykłady kryształów białka AnsA z warunków siarczanowych.



W przypadku kryształów, które rosły w warunkach zawierających $(NH_4)_2SO_4$ do optymalizacji zastosowałam komercyjny zestaw roztworów krystalizacyjnych Grid Screen Ammonium Sulfate (kod: HR2-211) firmy Hampton Research. Zawiera on 24 roztwory składające się z siarczanu amonu o stężeniach 0.8 M, 1.6 M, 2.4 M i 3.0 M, połączone z buforami o pH od 4.0 do 9.0. Kryształy jakie otrzymałam nadal były pozrastane, kruche i łatwo rozpadały się na drobne igły.

Optymalizację pozostałych dwóch warunków siarczanowych (zawierających Li_2SO_4) rozpoczęłam przygotowując dwa roztwory, po 100 mL, o składzie identycznym z roztworami komercyjnymi (Tabela 6.8). Oba rozcieńczałam wodą, aby wszystkie ich składniki stanowiły od 100 do 40% mieszaniny wyjściowej. Tak przygotowane roztwory służyły zarówno do nastawiania zwykłych krystalizacji białka AnsA, jak i szczepienia zarodkami. Stosowane stężenia białka wahały się od 5 do 20 mg/mL, natomiast proporcje składników w kropli wynosiły: 1:5, 2:3, 1:3, 2:2, 3:3, 3:1, 3:2 i 5:1 µL. Płytki z kroplami krystalizacyjnymi umieszczałam w temperaturze 4, 10 lub 19ºC.

Warunek nr 35 z zestawu Structure Screen 1, zawierający PEG 4K pozwalał uzyskać kryształy pokroju płytkowego, bez igieł, choć nadal mocno zrośnięte, dlatego kontynuowałam jego optymalizacje. Do mieszaniny stanowiącej 90, 80, 70 i 60% wyjściowego warunku dodawałam roztwory o następujących stężeniach końcowych:

- glicerol (3%, 6%, 12%),
- glicynę (0.1 M, 0.2 M, 0.4 M),
- oktyloglukozę (0.5%, 1.0%, 2.0%),
- dioksan (3%, 6%, 12%),
- PEG 400 (5%, 10%, 20%).

Tabela 6.9. Przykładowy zestaw roztworów rezerwuarowych służący optymalizacji warunku siarczanowego nr 35.

0.2 M Li_2SO_4 0.1 M HEPES pH 7.5 30% PEG 2K MME	0.2 M Li_2SO_4 0.1 M Tris pH 8.5 30% PEG 2K MME	0.2 M Li_2SO_4 0.1 M Gly/NaOH pH 9.5 30% PEG 2K MME
0.2 M Li_2SO_4 0.1 M HEPES pH 7.5 25% PEG 3350	0.2 M Li_2SO_4 0.1 M Tris pH 8.5 25% PEG 3350	0.2 M Li_2SO_4 0.1 M Gly/NaOH pH 9.5 25% PEG 3350
0.2 M Li_2SO_4 0.1 M HEPES pH 7.5 30% PEG 4K	0.2 M Li_2SO_4 0.1 M Tris pH 8.5 30% PEG 4K	0.2 M Li_2SO_4 0.1 M Gly/NaOH pH 9.5 30% PEG 4K
0.2 M Li_2SO_4 0.1 M HEPES pH 7.5 20% PEG 6K	0.2 M Li_2SO_4 0.1 M Tris pH 8.5 20% PEG 6K	0.2 M Li_2SO_4 0.1 M Gly/NaOH pH 9.5 20% PEG 6K

Przygotowywałam także nowe zestawy roztworów bazujące na warunku 35. Z trzech składników tego warunku jeden zawsze pozostawiałam niezmieniony, a dwa modyfikowałam (Tabela 6.9).

Pomimo optymalizacji warunków siarczanowych, podczas których nastawiłam w sumie ok. 500 kropel krystalizacyjnych, nie uzyskałam kryształów nadających się do pomiarów dyfrakcyjnych.

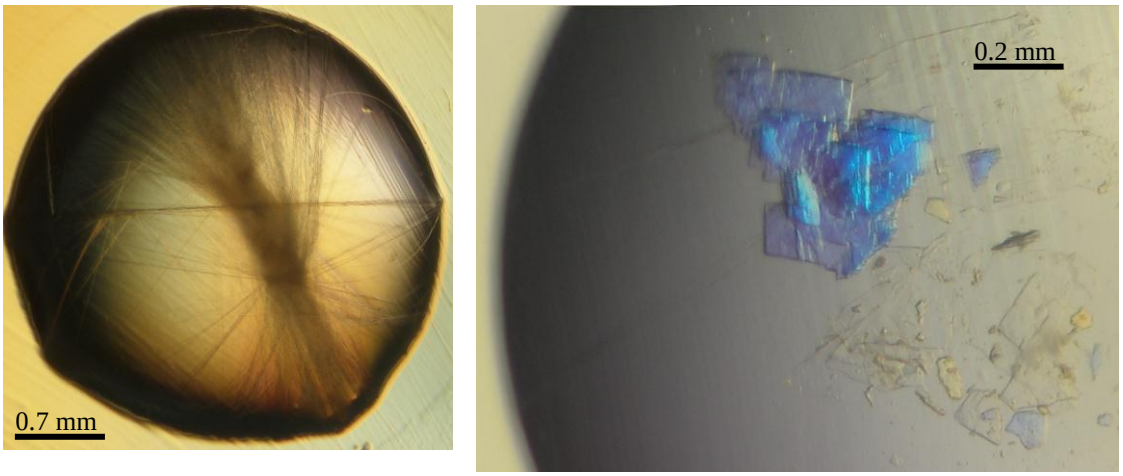
Optymalizacja warunków octanowych

Warunki zawierające octany (Tabela 6.10) optymalizowałam używając preparatów białka C, E i I. W tym przypadku kryształy również miały różny pokrój. Były igłami, które nie nadawały się nawet do testów rozpraszania (w dwóch warunkach zawierających CH_3COONa) lub rosły jako nieregularnie pozlepiane płytki słabo rozpraszające promieniowanie rentgenowskie (w warunku z $(CH_3COO)_2Ca$) (Rysunek 6.5)

Tabela 6.10. Warunki octanowe krystalizacji białka AnsA.

nazwa zestawu do krystalizacji (kod) producent numer warunku	skład roztworu
Clear Strategy Screen I (MD1-14) MDL 13	0.3 M CH_3COONa 10% PEG 8K 10% PEG 1K
Clear Strategy Screen I (MD1-14) MDL 19	0.3 M CH_3COONa 8% PEG 20K 8% PEG 550 MME
Clear Strategy Screen II (MD1-15) MDL 12	0.2 M $(CH_3COO)_2Ca$ 8% PEG 20K 8% PEG 550 MME

Rysunek 6.5. Przykłady kryształów białka AnsA z nieoptymalizowanych warunków octanowych.



Podczas optymalizacji warunków octanowych przygotowywałam roztwory zawierające CH_3COONa lub $(CH_3COO)_2Ca$ o różnych stężeniach, łączone z różnymi stężeniami i rodzajami glikoli polietylenowych (Tabela 6.11). Dodatkowo do roztworów tych dodawałam bufory modyfikując pH: cytrynianowy (0.1 M, pH 6.5), HEPES (0.1 M, pH 7.5) i Tris (0.1 M, pH 8.5). Podczas optymalizacji stosowałam temperaturę 19ºC, preparaty białka E i I o stężeniach 13 mg/mL i proporcje preparatu białka do roztworu rezerwuarowego: 1:3, 2:2, 3:1 µL.

Tabela 6.11. Początkowa optymalizacja warunków octanowych. Wymieniam zmienne elementy roztworów rezerwuarowych i te ich wartości, które były stosowane. Jednostki podaję w nawiasach.

CH_3COONa [M]	0.2 0.3 0.4
$(CH_3COO)_2Ca$ [M]	0.1 0.2 0.3
PEG 2K [%]	15 20 25 30
PEG 4K [%]	6 9 12 15
PEG 3350 [%]	12 16 20 25
PEG 1K / PEG 8K [% /%]	6/6 8/8 10/10 12/12
PEG 550 MME / PEG 20K [% /%]	4/4 6/6 8/8 12/12

Pomimo nastawienia ok. 500 kropli nie uzyskałam żadnych kryształów. Przy niższych stężeniach odczynników krople pozostawały czyste, natomiast przy wyższych białko strącało się. Dlatego przystępując do ponownych optymalizacji postanowiłam używać tylko oryginalnego komercyjnego roztworu, w którym uzyskiwałam kryształy pokroju płytkowego: warunku 12 z zestawu Clear Strategy Screen II (MD1-15) (MDL).

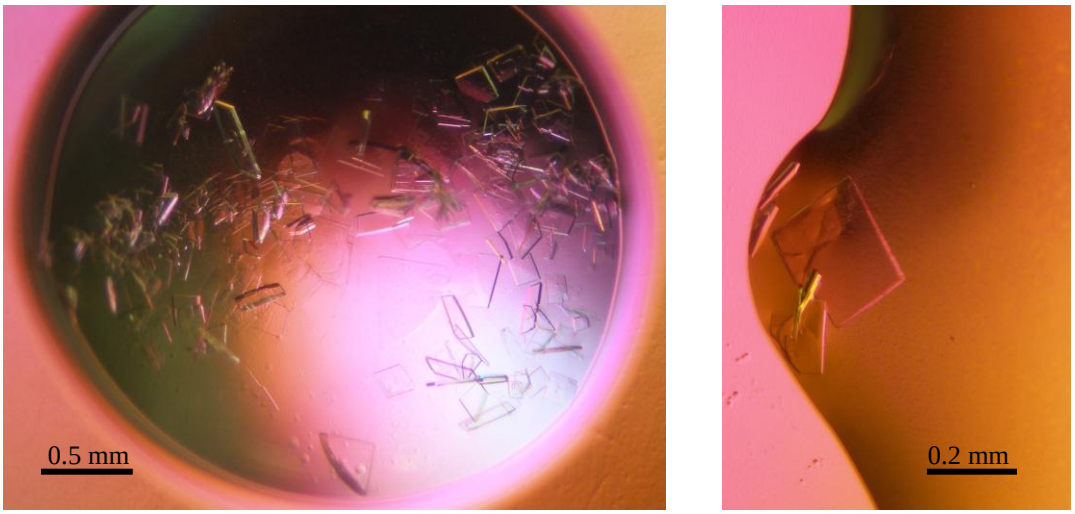
Ponownie nastawiłam ok. 500 kropli. Nie mogąc modyfikować stężeń poszczególnych składników komercyjnej mieszaniny rozcieńczałam ja przygotowując szereg nowych roztworów zawierających od 90 do 40% wyjściowej mieszaniny (Tabela 6.12). Roztwory łączyłam z różnymi odczynnikami o różnych stężeniach. Nastawiałam krople wiszące i siedzące w zmiennych proporcjach i objętościach używając preparatów białka C, E i I o różnych stężeniach. Próbowałam także szczepić krople zarodkami krystalizacyjnymi używając różnych technik wykonania tego zabiegu.

Ponowna optymalizacja warunku octanowego nr 12 z zestawu Clear Strategy Screen II (MD1-15) firmy MDL pozwoliła uzyskać kryształy w postaci oddzielnych płytek (Rysunek 6.6), które nadawały się do pomiarów dyfrakcji. Kryształy te rosły w kroplach zawierających preparat białka E (zawierający Tris i TCEP) o stężeniu 20 mg/mL w proporcji 1 µL białka do 3 - 5 µL rezerwuaru, który z kolei zawierał 60 - 80% oryginalnego komercyjnego roztworu i dodatek 1 - 2% oktyloglukozy.

Tabela 6.12. Ponowna optymalizacja warunku octanowego nr 12 z zestawu Clear Strategy Screen II. Wymieniam zmienne elementy i te ich wartości, które były stosowane. Jednostki podaję w nawiasach.

zawartość komercyjnego roztworu [%]	90 80 75 70 65 60 50 40
pH roztworów rezerwuarowych	6.5 (0.1 M bufor cytrynianowy) 7.5 (0.1 M HEPES) 8.5 (0.1 M Tris) 9.5 (0.1 M Gly/NaOH)
glicerol [%]	3 6 12
glicyna [M]	0.1 0.2 0.4 0.8
Arg / Glu [mM/mM]	25/25 50/50 100/100
oktyloglukoza [%]	0.5 1.0 1.5 2.0 3.0 4.0
dioksan [%]	3 6 12
PEG 400 [%]	5 10 20
TMAO [M]	0.05 0.10 0.25
NDSB [M]	0.05 0.10 0.25
MPD [%]	3 6 12
glikol etylenowy [%]	3 6 12
jeffamina M600 [%]	2 4 8
stężenia preparatu białka [mg/mL]	7 10 13 16 20
proporcje składników kropli [µL]	1:6 1:5 1:4 1:3 1:2 2:3 1:1 6:1 5:1 4:1 3:1 2:1 3:2 2:2
temperatura [°C]	4 19 30

Rysunek 6.6. Przykłady kryształów białka AnsA ze zoptymalizowanego warunku nr 12 z zestawu Clear Strategy Screen II (MD1-15) firmy MDL .



6.2.5.3. Próby uzyskania kryształów do badań dyfrakcyjnych

W związku z założeniem, że do rozwiązania struktury asparaginazy z *Rhizobium etli* najprawdopodobniej nie wystarczy metoda podstawienia cząsteczkowego, starałam

się również uzyskać kryształy umożliwiające zarejestrowanie sygnału anomalnego. Wypróbowałam dwa podejścia: nasączanie natywnych kryształów białka AnsA kompleksem $[Ta_6Br_{12}]Br_2$ oraz krystalizację białka znakowanego selenem.

Kryształy nasączane związkiem tantalu

Do nasączania wykorzystywałam kryształy, które hodowałam nastawiając wiele kropel zoptymalizowanego warunku octanowego. Roztwór $[Ta_6Br_{12}]Br_2$ dodawałam bezpośrednio do kropli z kryształami, przy czym stosowałam różne jego objętości (0.2 – 1.0 μL) i stężenia (5 - 100 mM) w zależności od wielkości kropli krystalizacyjnej i ilości/wielkości kryształów. Nasączanie prowadziłam w czasie od 1 do 16 godzin. Przygotowane kryształy były koloru zielonego.

Kryształy białka znakowanego selenem

Oczyszczone białko AnsA-Se dializowałam do roztworu zawierającego 20 mM Tris pH 7.5 i 5 mM TCEP, a następnie zatężałam do 16 mg/mL. Za pomocą robota krystalizacyjnego nastawiłam krople korzystając z ośmiu komercyjnych zestawów do krystalizacji. Poszukiwanie wstępnych warunków wykazało, że znakowane atomami selenu białko krystalizuje w tych samych warunkach co białko natywne. Nowy preparat również tworzył pozrastane kryształy w warunkach fosforanowych, siarczanowych oraz octanowych. Posłużyłam się zatem oktyloglukozą, dzięki której poprawiłam pokrój i zdolności dyfrakcyjne kryształów białka natywnego i w tak zoptymalizowanym warunku octanowym uzyskałam kryształy AnsA-Se w postaci pojedynczych płytek. Mimo podobnej morfologii znakowane kryształy różniły się od natywnych. Przede wszystkim były znacznie mniejsze, a podczas wyławiania ich z kropli zauważyłam, że są elastyczne i trudno je przełamać, co jest raczej bardzo rzadko spotykane.

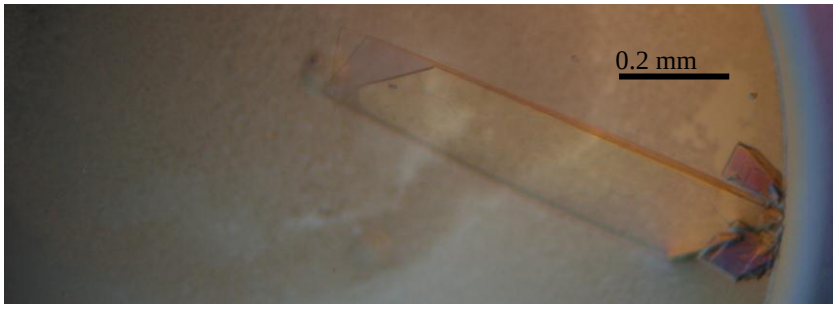
6.2.6. Pomiary dyfrakcji rentgenowskiej

Uzyskane przeze mnie kryształy były poddawane testom rozpraszania promieniowania rentgenowskiego na synchrotronie BESSY w Berlinie na linii pomiarowej 14.2. Tam też najlepsze kryształy służyły do rejestracji danych dyfrakcyjnych. Testy i pomiary kryształów wykonywałam osobiście lub korzystałam z

pomocy kolegów z CBB, gdy nie miałam możliwości wyjazdu. Wszystkie pomiary przeprowadzono w temperaturze 100 K zapewnionej przez strumień par azotu. Przetestowane krioprotektanty (np. glicerol, PEG400) miały zły wpływ na jakość rozpraszania, dlatego kryształy przygotowywane do pomiarów były osuszane, a następnie mrożone w ciekłym azocie.

Spośród wszystkich kryształów natywnego białka AnsA uzyskanych ze zoptymalizowanego warunku octanowego wybrano jeden, który najlepiej rozpraszał promieniowanie rentgenowskie. Kryształ ten (Rysunek 6.7), a właściwie jego fragment (Tabela 6.13, kryształ P3) pozwolił na rejestrację zbioru danych o rozdzielczości 1.5 Å. Podczas pomiarów zastosowano standardową długość fali (0.91841 Å) przy której, na danej linii pomiarowej, występuje maksimum intensywności promieniowania.

Rysunek 6.7. Kryształ natywnego białka AnsA rozpraszający promieniowanie rentgenowskie do 1.5 Å.



Podjęto również próby rejestracji sygnału anomalnego. Niestety, elastyczne kryształy znakowanego białka AnsA-Se w ogóle nie rozpraszały promieniowania rentgenowskiego. Z kolei kryształy natywnego białka nasączone kompleksem tantalu rozpraszały znacznie gorzej od nienasączonych. Mimo to, jeden z kryształów rozpraszający do rozdzielczości 3 Å posłużył do rejestracji zbioru obrazów przy długości fali (1.25 Å) wybranej na podstawie krzywej fluorescencji (Tabela 6.13, kryształ 16T). Kryształ ten podczas naświetlania uległ zniszczeniom radiacyjnym, dlatego od ok. 650 obrazu refleksy zaczęły zanikać.

Przygotowane przeze mnie kryształy zawierające wprowadzone różnymi metodami atomy specjalne Se i Ta nie pozwoliły na uzyskanie zadowalającego wyniku. Postanowiłam zatem podjąć próbę rejestracji sygnału anomalnego pochodzącego od naturalnie występujących w białku atomów siarki, ponieważ nie wymagała ona ingerencji w stosunkowo dobrze rozpraszające kryształy natywnego białka AnsA. Monomer AnsA (367 reszt aminokwasowych) zawiera aż 21 atomów siarki w łańcuchach bocznych cysteiny i metioniny.

Nowo przygotowane kryształy natywnego białka AnsA ze zoptymalizowanego warunku octanowego rozpraszały promieniowanie rentgenowskie do rozdzielczości 2.0 Å. Najlepsze z nich wykorzystano do rejestracji zbiorów danych dyfrakcyjnych przy długości fali 1.9 i 2.0 Å (Tabela 6.13), umożliwiających detekcję słabego sygnału anomalnego od atomów siarki. Każdy kryształ poddawany był dwu- lub trzykrotnemu naświetlaniu w różnej orientacji, aby zapewnić wystarczająco dużą redundancję niezbędną do rozwiązania struktury metodą SAD. To z kolei powodowało, że kryształy ulegały zniszczeniu i przestawały rozpraszać, jak w przypadku kryształów 9asp i 1asp, dla których rejestracja trzeciego zestawu danych została przerwana (Tabela 6.13).

Tabela 6.13. Zarejestrowane zbiory danych dyfrakcyjnych dla kryształów białka AnsA. Podałam podstawowe parametry pomiarów z jednostkami w nawiasach, * oznacza zmianę ustawienia kryształu względem wiązki, która uzyskana była przez umyślne zgięcie pętelki zawierającej kryształ.

nazwa kryształu	numer zbioru	długość fali [Å]	odległość detektora [mm]	początek rejestracji [º]	oscylacja [º]	czas naświetlania [s]	liczba obrazów
P3	Nat - P3 - 7	0.91841	180	0	0.5	10	400
16T	Ta - 16T - 4	1.255	202	30	0.5	5	720
9asp	S - 9asp - 5	1.9	130	0	0.5	7	720
	S - 9asp - 6	1.9	135	0.3	0.5	7	720
	S - 9asp - 7	1.9	140	0.6	0.5	7	210
1asp	S - 1asp - 5	1.9	100	0	0.5	5	720
	S - 1asp - 6	1.9	95	0.25	0.5	5	720
	S - 1asp - 7	1.9	110	*	0.5	5	322
7asp	S - 7asp - 2	1.9	100	0	0.5	3	720
	S - 7asp - 3	1.9	110	*	0.25	3	720
	S - 7asp - 4	1.9	110	**	0.25	3	720
A9	S - A9 - z1	2.0	90	0	0.5	5	720
	S - A9 - z2	2.0	95	0.25	0.5	5	720
A5	S - A5 - 3	2.0	100	0	1	8	360
	S - A5 - 4	2.0	100	0.5	1	8	360

6.2.7. Weryfikacja danych dyfrakcyjnych

Wszystkie zarejestrowane zbiory danych dyfrakcyjnych (Tabela 6.13) podlegały procesowaniu, które umożliwiało ich weryfikację. Przetwarzanie danych dyfrakcyjnych wykonałam za pomocą programu *XDS* (**Kabsch, 2010**). W trakcie procesowania ustalałam limit rozdzielczości tak, aby uzyskać jak najwyższą kompletność danych i aby stosunek sygnału do szumu $\langle I/\sigma I \rangle$ w ostatnim przedziale rozdzielczości był bliski 2. Prawie wszystkie zbiory obrazów dyfrakcyjnych (za wyjątkiem „natywnego" Nat - P3 - 7) procesowałam z założeniem „Friedel's law = FALSE", co umożliwia detekcję

sygnału anomalnego. Nie wszystkie zbiory danych dyfrakcyjnych udało się przeprocesować, dlatego prezentuję tylko te, dla których wskaźnikowanie, integracja i skalowanie powiodły się (Tabela 6.14).

Tabela 6.14. Wyniki przetwarzania zbiorów danych dyfrakcyjnych, których procesowanie powiodło się. W nawiasach kwadratowych podano jednostki, a w okrągłych wartości dotyczące ostatniego zakresu rozdzielczości.

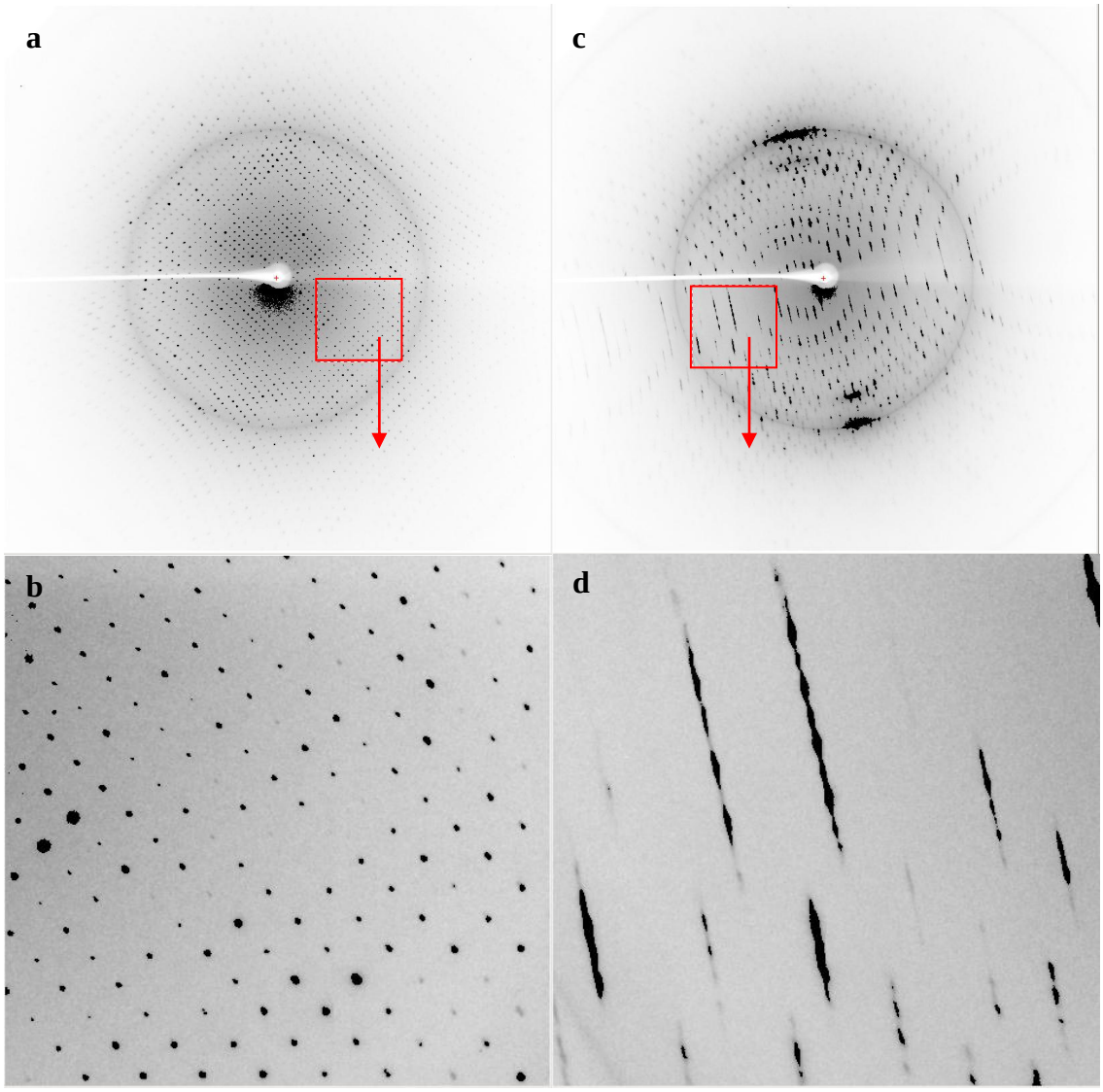
numer zbioru	parametry komórki [Å] | [º]		grupa przestrzenna	rozdzielczość [Å]	mozaikowość [º]	krotność	kompletność [%]	R_{merge} [%]	I/σI [%]	sygnał anomalny
Nat - P3 - 7	77.98 91.46 114.19	90.00 96.89 90.00	$P2_1$	1.46	0.197	3.72	91.2 (60.4)	8.1 (50.0)	10.66 (1.98)	nie dotyczy
Ta - 16T - 4	78.06 120.92 91.05	90.00 90.11 90.00	$P2_1$	2.90	0.151	3.82	98.5 (97.1)	13.5 (48.9)	9.81 (3.07)	słaby od 5.0 Å
S - 1asp - 5	77.63 91.27 113.91	90.00 97.26 90.00	$P2_1$	2.23	0.429	3.58	98.2 (90.6)	8.4 (26.8)	11.54 (4.36)	słaby od 4.7 Å
S - 7asp - 2	77.81 91.28 114.30	90.00 90.00 90.00	$P222$	2.23	0.350	7.09	97.3 (83.1)	10.2 (22.1)	15.06 (7.00)	słaby od 6.5 Å
S - 7asp - 3	77.86 91.37 114.32	90.00 90.00 90.00	$P222_1$	2.34	0.376	3.53	97.1 (82.2)	8.6 (17.0)	11.14 (4.58)	słaby od 7.0 Å
S - 7asp - 4	77.81 91.39 114.22	90.00 90.00 90.00	$P222$	2.34	0.393	3.36	96.0 (79.7)	9.9 (26.6)	9.87 (3.42)	brak
S - A9 - z1	78.08 123.13 91.46	90.00 90.17 90.00	$P2$	2.45	0.466	3.52	93.4 (71.8)	12.0 (32.6)	7.63 (2.92)	słaby od 3.3 Å
S - A9 - z2	77.87 122.13 91.29	90.00 90.08 90.00	$P2$	2.51	0.517	2.99	92.4 (92.7)	9.7 (29.7)	8.63 (3.26)	słaby od 2.5 Å
S - A5 - 3	78.26 91.77 124.88	90.00 90.00 90.00	$P222$	2.23	0.481	6.56	94.3 (70.0)	15.8 (33.6)	9.28 (2.70)	brak
S - A5 - 4	78.29 91.92 123.18	90.00 90.00 90.00	$P222$	2.36	0.481	6.94	98.6 (91.3)	16.6 (38.2)	9.12 (3.51)	słaby od 7.0 Å

Niestety dane uzyskane w wyniku dyfrakcji promieniowania o długości fali leżącej w pobliżu progu absorpcji atomów siarki nie nadają się do rozwiązania struktury metodą SAD, ze względu na brak wystarczająco silnego sygnału anomalnego. Przeprocesowane dane pochodzące ze zbioru obrazów Ta - 16T - 4 także nie nadają się do dalszych obliczeń z tego samego powodu. Jedynie zbiór obrazów dyfrakcyjnych Nat

- P3 - 7, zarejestrowanych dla najlepiej rozpraszającego kryształu natywnego białka AnsA (P3) przy użyciu promieniowania o maksymalnej intensywności (λ = 0.91841 Å) udało mi się przetworzyć w zbiór refleksów, który kwalifikuje się do dalszej obróbki.

Co ciekawe, przeprocesowane dane wykazują różne grupy przestrzenne o zbliżonych stałych sieciowych. Aby skontrolować obecność osi śrubowych przeprowadziłam analizę wygaszeń systematycznych, jednak wynik nie był do końca jednoznaczny. Nie udało mi się więc stwierdzić, czy różne grupy przestrzenne poszczególnych zestawów danych wynikają z faktycznych różnic w symetrii kryształów, czy jest to jedynie artefakt spowodowany złą jakością danych.

Rysunek 6.8. Przykładowe obrazy dyfrakcyjne zarejestrowane dla kryształu A5, pochodzące z tego samego zbioru S - A5 - 3: obrazy nr 210 (**a**) oraz nr 120 (**c**). Narożnik detektora odpowiada rozdzielczości 2.1 Å. Zaznaczone czerwoną linią fragmenty obrazów, prezentuję na powiększeniach (**b** i **d**). Obraz (**c**) został zarejestrowany po obrocie kryształu o kąt φ=90°, w stosunku do obrazu (**a**).



W związku z przedstawionymi powyżej problemami, wszystkie zgromadzone dla białka AnsA zbiory obrazów dyfrakcyjnych poddałam również ocenie wizualnej. Analizę tą przeprowadziłam w programie Adxv (**Adxv, 2013**), który umożliwia szybkie wyświetlenie kolejnych dwuwymiarowych obrazów zarejestrowanych podczas danego eksperymentu dyfrakcyjnego. Dzięki temu zauważyłam, że wszystkie kryształy (nasączane związkiem tantalu oraz natywne) rozpraszały promieniowanie rentgenowskie anizotropowo, tj. następowała zmiana w jakości rozpraszania w zależności od orientacji kryształu. W każdym ze zbiorów w pewnych kierunkach obserwowałam refleksy, które nie były wyraźnie rozdzielone, lecz łączyły się w smugi (Rysunek 6.8). Taki efekt „rozmytych" refleksów wynikać może z częściowego nieuporządkowania kryształu w tym kierunku lub ze zbliźniaczenia, jednak dalszych analiz już nie przeprowadzałam.

6.2.8. Ocena jakości preparatu białka

Oczyszczone preparaty białka AnsA i AnsA-Se przed krystalizacją poddawane były ocenie jakościowej. Weryfikowałam różne cechy białka różnymi metodami.

6.2.8.1. Test aktywności enzymatycznej

Charakterystyka biochemiczna białka AnsA jest dobrze znana i opublikowana (**Moreno-Enriquez *et al*., 2012**). Celem testu enzymatycznego jaki zaplanowałam dla białka AnsA było sprawdzenie, czy uzyskany przeze mnie preparat jest aktywny w temperaturze krystalizacji, tj. 19ºC. Test polegał na reakcji hydrolizy asparaginy i śledzeniu jej postępu w czasie za pomocą detekcji powstających jonów amonowych. Do kolorymetrycznego oznaczenia stężenia jonów NH_4^+ stosowałam odczynnik Nesslera. Jest to zasadowy roztwór tetrajodortęcianu (II) potasu $K_2[HgI_4]$, który wchodząc w reakcję z jonami NH_4^+ tworzy związek kompleksowy NH_2Hg_2OI o zabarwieniu żółto-pomarańczowym:

$$NH_4Cl + 2\ K_2[HgI_4] + 4\ KOH \rightarrow NH_2Hg_2OI + KCl + 7\ KI + 3\ H_2O$$

Podstawą przeprowadzonego przeze mnie testu aktywności enzymatycznej były eksperymenty, jakie opisała dr Karolina Michalska w swojej rozprawie doktorskiej (**Michalska, 2007**). Przygotowanie testu wymagało kilku wstępnych eksperymentów, podczas których dobrałam odpowiednie stężenia substratu (L-Asn) i enzymu, oraz

proporcje wszystkich odczynników, tak aby wyniki pomiarów kolorymetrycznych znajdowały się w wiarygodnym zakresie.

Dopracowany test polegał na prowadzeniu reakcji w objętości 1 mL, w buforze o pH 7.5 (20 mM Tris), w standardowej temperaturze krystalizacji (19ºC). Mieszanina reakcyjna zawierała L-Asn o końcowym stężeniu 10 mM i 5 μg enzymu AnsA. W dziesięciominutowych odstępach czasu pobierałam próbkę mieszaniny (50 μL) do której dodawałam 100 μL 15% kwasu trichlorooctowego aby zatrzymać reakcję. Następnie mieszaninę reakcyjną rozdzielałam na trzy równe porcje po 40 μL i do każdej dodawałam 760 μL odczynnika Nesslera rozcieńczonego wodą w stosunku 1:6.5. Absorbancję roztworów mierzyłam po 20 minutach przy długości fali λ= 414 nm. Trzy pomiary dla rozdzielonej próbki uśredniałam. Stężenie jonów NH_4^+ odczytywałam z krzywej kalibracyjnej przygotowanej dla roztworów siarczanu amonu. Jako próbę ślepą stosowałam analogiczną mieszaninę bez dodatku enzymu. Wyniki testu pokazują, że białko AnsA jest aktywne, ponieważ wraz z upływem czasu reakcji stężenie jonów amonowych rośnie (Tabela 6.15).

Tabela 6.15. Wyniki testu aktywności enzymatycznej enzymu AnsA.

czas reakcji [min.]	10	20	30	40	50	60	70	80	90	100	110
stężenie jonów NH_4^+ [μM]	56	92	114	130	136	142	144	148	150	150	152

6.2.8.2. Dynamiczne rozpraszanie światła

Białko AnsA poddawałam również testom na analizatorze Zetasizer Nano (Malvern Instruments Ltd.), który wykorzystuje zjawisko dynamicznego rozpraszania światła do określania stopnia dyspersji i promienia hydrodynamicznego cząsteczek zawartych w roztworze. Dzięki pomiarom DLS mogłam stwierdzić, czy otrzymywane przeze mnie preparaty białka są homogenne.

Za pomocą DLS badałam wszystkie preparaty AnsA oraz AnsA-Se, które poddawałam krystalizacji. Preparaty te różniły się składem: zawierały różne bufory (Tris, fosforanowy), różne czynniki redukujące (β-merkaptoetanol, DTT, TCEP), różne dodatkowe odczynniki (NaCl, L-Asp, glicerol) i różne stężenie białka. Pomiary wykonywałam dla preparatów świeżo oczyszczonych, przechowywanych w 4ºC przez różny okres czasu (niekiedy do miesiąca) oraz rozmrożonych po przechowywaniu w -80ºC.

Dzięki pomiarom DLS mogłam weryfikować, czy nowo przygotowane preparaty

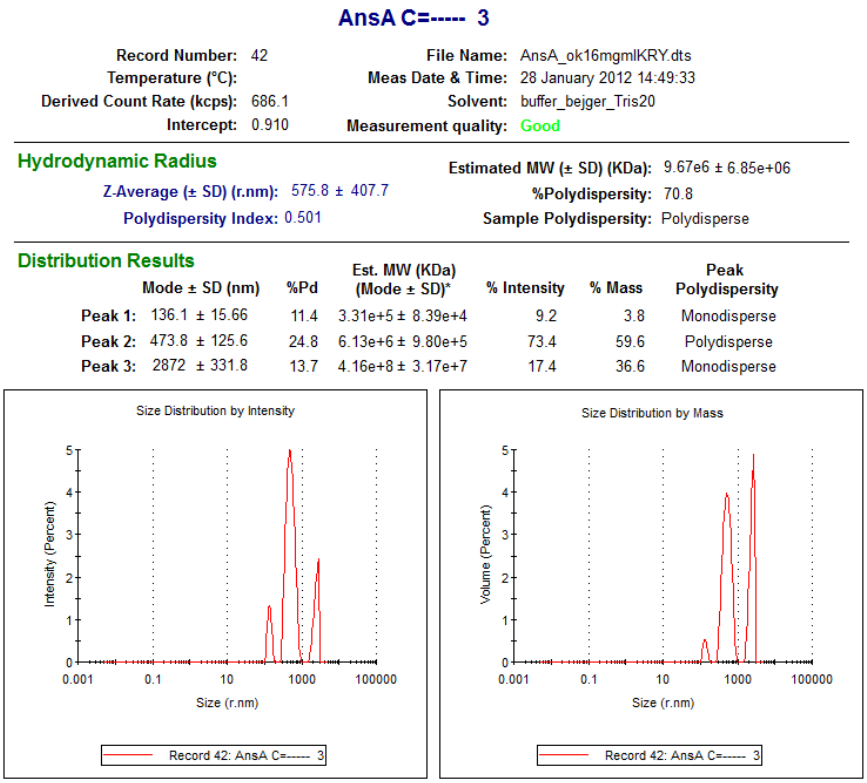
różnią się od wcześniejszych i czy ewentualne różnice wpływają w jakiś sposób na wyniki krystalizacji, lub też czy przechowywany preparat pozostał niezmieniony i można oczekiwać podobnych efektów krystalizacji.

Dzieląc wszystkie preparaty według wyników pomiarów DLS na dwie kategorie: dobre - wskazujące na obecność rozpuszczalnego białka oraz negatywne - ze względu na obecność agregatów, mogłam określić czynniki powodujące destabilizację białka AnsA/AnsA-Se. Jednym z nich jest brak czynnika redukującego (Rysunek 6.9), a drugim czas przechowywania preparatu w 4ºC (Rysunek 6.10). Najdłużej, nawet miesiąc, w niezmienionym stanie pozostawały preparaty o niższym stężeniu białka (do 10 mg/mL) jednocześnie zawierające TCEP, podczas gdy pozostałe powoli agregowały.

Biorąc pod uwagę jedynie preparaty zawierające rozpuszczalne białko wyniki pomiarów DLS były bardzo podobne. Badane próbki zawierały pulę jednorodnych cząsteczek (nie agregatów) o czym świadczył jeden monodyspersyjny pik na wykresach rozkładu wielkości cząsteczek (Rysunki 6.11, 6.12, 6.13, 6.14). Zatem preparaty te były homogenne. Promień obserwowanych cząsteczek najczęściej wynosił ok. 3.5 nm (Rysunek 6.12), choć zdarzały się również tak skrajne wyniki jak ok. 3.2 nm (Rysunek 6.14), czy ok. 3.9 nm (Rysunek 6.11). Rozbieżności są głównie efektem błędu pomiarowego wynikającego z różnych lepkości preparatów, choć sam błąd obliczeniowy również ma spore znaczenie. Dla przykładu, pokazuję wyniki dwóch identycznie przeprowadzonych pomiarów tej samej próbki (Rysunki 6.12 i 6.13). Co więcej, nawet podczas jednego pomiaru analizator podaje dwa (czasem znacznie różniące się) wyniki promienia cząsteczek liczone dwoma różnymi algorytmami (Rysunki 6.11 i 6.13). Oszacowana masa białka jest obarczona jeszcze większym błędem, ponieważ liczona jest z promienia dla obiektu idealnie sferycznego. Tak więc uzyskany zakres 52 - 80 kDa świadczy jedynie o tym, że kształt molekuł rozpuszczalnych białek AnsA/AnsA-Se w roztworze odbiega od sferycznego.

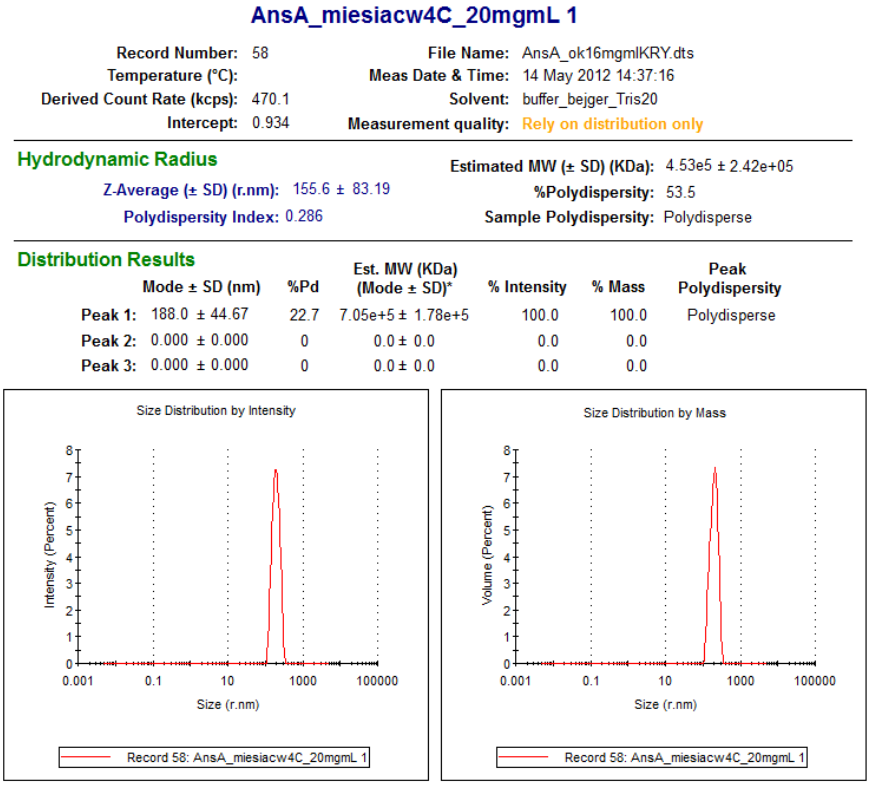
Przygotowałam również kilka testów, w których wykorzystywałam DLS do analizy wpływu dodatkowych czynników fizyko-chemicznych na homogenność preparatu białka AnsA. Testowałam wpływ temperatury wykonując pomiary w 15, 18, 24 i 30ºC dla tych samych próbek. Porównywałam wyniki DLS dla specjalnie przygotowanych preparatów białka różniących się jedynie wartością pH w zakresie 6.0 – 9.0, czy też zawierających różne kationy często występujące w warunkach, krystalizacyjnych (dodatki 5 mM Ca_2Cl, KCl, NaCl, LiCl lub Mg_2Cl). Żadna z tych zmiennych nie miała znaczącego wpływu na wyniki pomiarów DLS.

Rysunek 6.9. Przykładowy wynik pomiaru DLS dla zagregowanego preparatu AnsA, który nie zawierał czynnika redukującego (preparat F: 13 mg/mL białka AnsA, 20 mM Tris, pH 7.5).



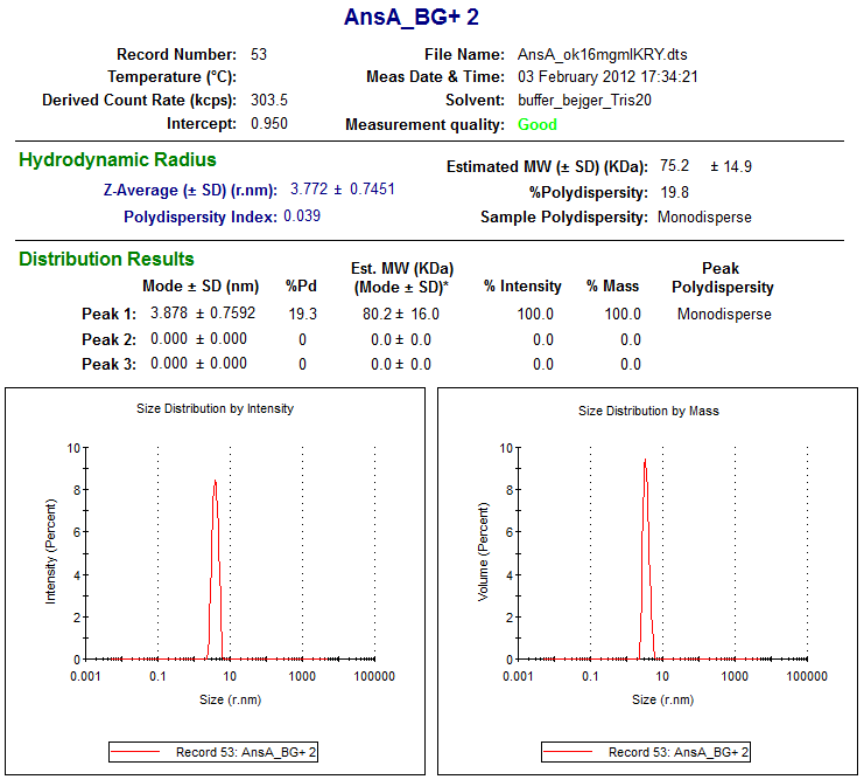
*The molecular weight reported here is only an estimate, calculated using an empirical mass vs. size calibration curve.

Rysunek 6.10. Przykładowy wynik pomiaru DLS zagregowanego preparatu AnsA, który był przechowywany w 4ºC przez miesiąc. Preparat zawierał: białko AnsA 20 mg/mL, 20 mM Tris pH 7.5, 5 mM TCEP.



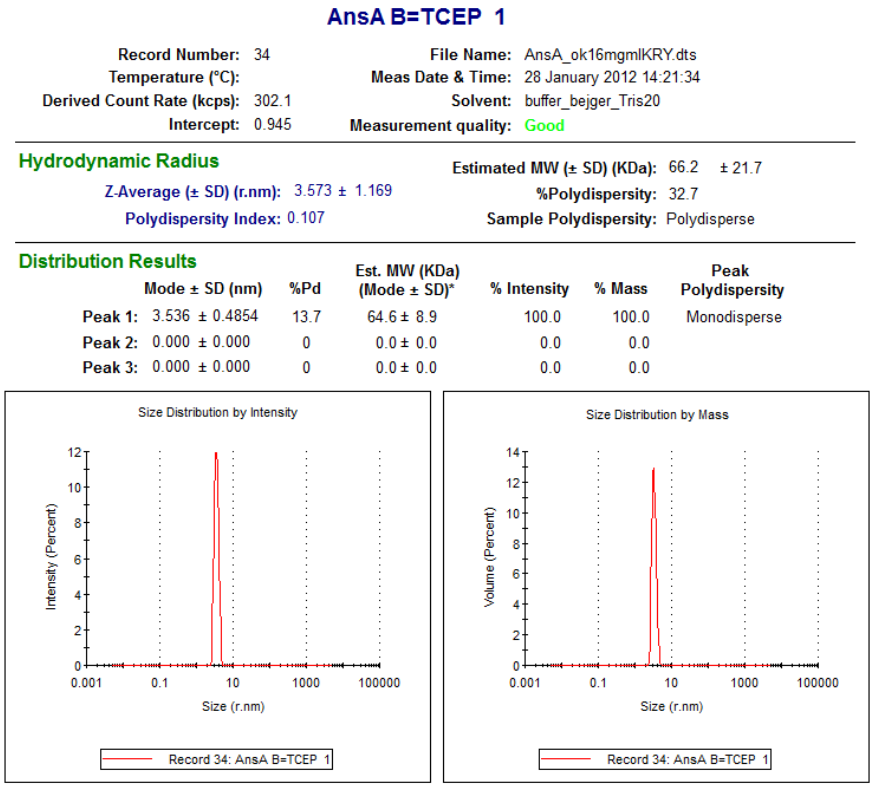
*The molecular weight reported here is only an estimate, calculated using an empirical mass vs. size calibration curve.

Rysunek 6.11. Przykładowy wynik pomiaru DLS dla preparatu AnsA, który zawierał glicerol (preparat G: 13 mg/mL białka AnsA, 20 mM Tris, pH 7.5, 5 mM TCEP, 5% glicerol).



*The molecular weight reported here is only an estimate, calculated using an empirical mass vs. size calibration curve.

Rysunek 6.12. Przykładowy wynik pierwszego pomiaru DLS dla preparatu AnsA, w którym składnikami roztworu były: białko AnsA 13 mg/mL, 20 mM Tris, pH 7.5, 5 mM TCEP (preparat E).



*The molecular weight reported here is only an estimate, calculated using an empirical mass vs. size calibration curve.

Rysunek 6.13. Przykładowy wynik drugiego pomiaru DLS dla preparatu AnsA, w którym składnikami roztworu były: białko AnsA 13 mg/mL, 20 mM Tris, pH 7.5, 5 mM TCEP (preparat E).

AnsA B=TCEP 2

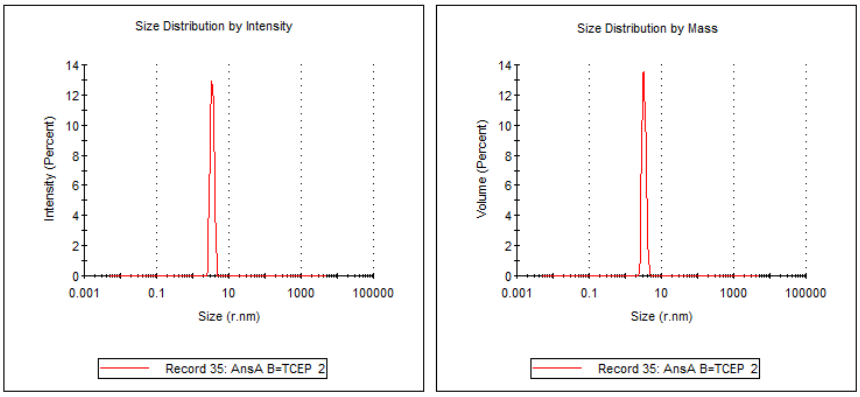
Record Number:	35	**File Name:**	AnsA_ok16mgmlKRY.dts
Temperature (°C):		**Meas Date & Time:**	28 January 2012 14:24:02
Derived Count Rate (kcps):	302.4	**Solvent:**	buffer_bejger_Tris20
Intercept:	0.946	**Measurement quality:**	Good

Hydrodynamic Radius

Z-Average (± SD) (r.nm): 3.517 ± 1.189
Polydispersity Index: 0.114
Estimated MW (± SD) (KDa): 63.9 ± 21.6
%Polydispersity: 33.8
Sample Polydispersity: Polydisperse

Distribution Results

	Mode ± SD (nm)	%Pd	Est. MW (KDa) (Mode ± SD)*	% Intensity	% Mass	Peak Polydispersity
Peak 1:	3.376 ± 0.4440	12.8	58.0 ± 7.9	100.0	100.0	Monodisperse
Peak 2:	0.000 ± 0.000	0	0.0 ± 0.0	0.0	0.0	
Peak 3:	0.000 ± 0.000	0	0.0 ± 0.0	0.0	0.0	



*The molecular weight reported here is only an estimate, calculated using an empirical mass vs. size calibration curve.

Rysunek 6.14. Przykładowy wynik pomiaru DLS dla preparatu AnsA-Se. Preparat zawierał: białko AnsA-Se 16 mg/mL, 20 mM Tris, pH 7.5, 5 mM TCEP.

AnsA_Se_rozmrozona_15mgmL 3

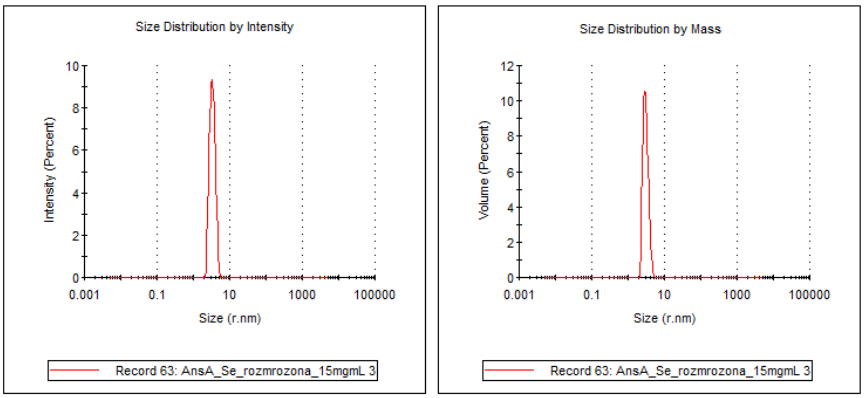
Record Number:	63	**File Name:**	AnsA_dojrzala.dts
Temperature (°C):		**Meas Date & Time:**	29 April 2014 17:00:44
Derived Count Rate (kcps):	309.4	**Solvent:**	buffer_bejger
Intercept:	0.890	**Measurement quality:**	Good

Hydrodynamic Radius

Z-Average (± SD) (r.nm): 3.278 ± 0.8191
Polydispersity Index: 0.062
Estimated MW (± SD) (KDa): 54.2 ± 13.5
%Polydispersity: 25.0
Sample Polydispersity: Polydisperse

Distribution Results

	Mode ± SD (nm)	%Pd	Est. MW (KDa) (Mode ± SD)*	% Intensity	% Mass	Peak Polydispersity
Peak 1:	3.224 ± 0.5992	17.8	52.1 ± 10.3	100.0	100.0	Monodisperse
Peak 2:	0.000 ± 0.000	0	0.0 ± 0.0	0.0	0.0	
Peak 3:	0.000 ± 0.000	0	0.0 ± 0.0	0.0	0.0	



*The molecular weight reported here is only an estimate, calculated using an empirical mass vs. size calibration curve.

6.2.8.3. Spektrometria mas

Preparaty białka AnsA zostały poddane analizie w spektrometrze masowym MALDI-TOF/TOF (UltrafleXtreme, Bruker) przeprowadzonej w Pracowni Spektrometrii Mass IChB PAN. Tożsamość oczyszczonego enzymu AnsA potwierdzono stosując analityczną technikę PMF (*ang. Peptide Mass Fingerprinting*), czyli peptydowego odcisku palca, w której widmo masowe peptydów uzyskanych po trawieniu enzymami proteolitycznymi porównuje się z widmami znajdującymi się w bazie danych. Spektrometria mas wykazała ponadto, że monomer AnsA ma masę 39588.378 Da co odpowiada teoretycznie wyliczonej w programie ProtParam masie rekombinowanego białka AnsA (39.5 kDa) uwzględniającej sześć aminokwasów pozostających na N-końcu po odcięciu znacznika histydynowego. Z kolei masa wyznaczona dla monomeru AnsA-Se wynosiła 40108.326 Da, co oznacza, że ok. jedenaście (z trzynastu obecnych w monomerze) reszt metioninowych zostało podstawionych Se-Met.

6.3. Dyskusja wyników

6.3.1. Weryfikacja wydajności procesów produkcji i oczyszczania białka AnsA

Optymalizację procesu produkcji białka AnsA przeprowadziłam bez większych problemów i szybko uzyskałam zadowalające rezultaty. Ekspresję natywnego białka AnsA prowadziłam z plazmidowego wektora ekspresyjnego pET-Re-AnsA w komórkach BL21Star, w pożywce TB, w temperaturze 25ºC, indukując ją 0.2 mM IPTG. Ekspresję białka AnsA-Se prowadziłam w bardzo podobnych warunkach, indukując ją 0.1 mM IPTG i stosując pożywkę minimalną, w której metioninę zastąpiłam selenometioniną. Ustalone warunki ekspresji zapewniały produkcję rozpuszczalnych białek w ilościach pozwalających na wydajne oczyszczanie, dlatego nie kontynuowałam optymalizacji tego etapu.

Opracowany przeze mnie ostateczny protokół oczyszczania (Tabela 6.4) umożliwia uzyskiwanie 14 mg czystego białka AnsA z 4 g osadu komórkowego i 10 mg czystego białka AnsA-Se również z 4 g osadu komórkowego. Procedura składa się z etapów: liza, chromatografia powinowactwa, cięcie proteazą TEV^{His}, chromatografia powinowactwa, chromatografia żelowa, wymiana buforu. Poszczególne etapy zostały

przeze mnie dopracowane co do ilości zużywanych odczynników, ich proporcji względem oczyszczanego preparatu oraz czasu trwania konkretnych kroków oczyszczania.

Początkowo starałam się, aby protokół oczyszczania zawierał jedynie niezbędne etapy i był na tyle krótki, aby umożliwić uzyskanie czystego białka w dwa dni. Ostateczny protokół zawiera dwa dodatkowe etapy, co wydłuża początkową procedurę do trzech dni. Jednak białka AnsA i AnsA-Se są wyjątkowo stabilne, niewrażliwe na zmiany temperatury i stężenia odczynników. Nieznacznie strącają się jedynie podczas zatężania. Cecha ta pozwala pracować bez pośpiechu i obawy, że przedłużona procedura spowoduje znaczne straty AnsA i AnsA-Se.

Położyłam zatem nacisk na wydajność protokołu i jakość uzyskiwanego preparatu kosztem czasu poświęconego na poszczególne etapy. Wydłużyłam proces lizy, czas inkubacji białek ze złożem niklowym oraz zwiększyłam ilość roztworu wykorzystywanego do przemywania białek związanych do złoża. Dodany etap chromatografii żelowej służy „doszlifowaniu" czystości preparatu (ang. *polishing step*).

W trakcie całej procedury stosuję zasadniczo cztery odczynniki: Tris, NaCl, β-merkaptoetanol i imidazol, natomiast na jej końcu dializuję oczyszczone białko do roztworu zawierającego jedynie Tris i TCEP. Mogłabym skrócić protokół zmieniając skład roztworu już podczas chromatografii żelowej, jednak usunięcie soli mogłoby niekorzystnie wpłynąć na prowadzony rozdział białek. Ponadto podczas dializy (prowadzonej na zatężonym preparacie, zatem w małych objętościach) zużywam mniej czynnika redukującego niż podczas chromatografii, co prowadzi do oszczędności drogiego odczynnika - TCEP. Wyodrębnienie tego ostatniego etapu - zmiany składu roztworu białka ma jeszcze jedną zaletę. Nie zmieniając zasadniczej procedury oczyszczania można przygotować preparat białka w dowolnych roztworach, co przydało się podczas poszukiwania wstępnych warunków krystalizacji.

6.3.2. Ocena jakości preparatu białka

Wyprodukowane i oczyszczone metodami biotechnologicznymi białka AnsA i AnsA-Se są preparatami dobrej jakości, homogennymi, charakteryzującymi się znaczną stabilnością potwierdzoną wynikami pomiarów DLS. Maksymalne stężenia jakie można uzyskać zatężając białka AnsA i AnsA-Se to odpowiednio 30 i 25 mg/mL. Potwierdziłam również, że białko AnsA jest aktywne w warunkach krystalizacyjnych.

Za pomocą spektrometrii mas zweryfikowałam tożsamość oczyszczonych białek, jak również obecność atomów selenu w białku AnsA-Se.

Oba białka są rozpuszczalne w dość różnorodnych roztworach, tolerują pH w zakresie 6 - 9 oraz zupełny brak soli. Warunkiem stabilności obu preparatów jest jednak obecność czynnika redukującego, ponieważ jego brak powoduje strącanie się białka, co widać na filtrach podczas zatężania, w kroplach krystalizacyjnych oraz w wynikach pomiarów DLS. Kluczowym składnikiem preparatów był TCEP, który znacząco i na długo poprawiał rozpuszczalność białek.

Aby zabezpieczyć białka, najczęściej mroziłam je w -80ºC w małych 50 µL porcjach, ponieważ ten sam preparat przed mrożeniem i po rozmrożeniu tak samo krystalizował oraz dawał bardzo zbliżone wyniki pomiarów DLS. Możliwe było również przechowywanie preparatów w 4ºC, przy czym im wyższe było stężenie białka, tym szybciej strącało się ono w miarę upływu czasu.

6.3.3. Analiza prób krystalizacyjnych

Podsumowując etap krystalizacji mogę stwierdzić, że zarówno natywne białko AnsA jak i znakowane białko AnsA-Se chętnie i dość szybko (po 1 - 14 dniach) krystalizuje w warunkach zawierających fosforany, siarczany i octany, przy czym jakość dyfrakcji na tych kryształach jest bardzo zróżnicowana. Warunkiem koniecznym do uzyskania kryształów jest odpowiedni skład preparatu białka, który wyklucza zawartość soli. Ważna jest też obecność czynnika redukującego, powstrzymującego przedwczesne strącanie się białka.

Znalezione podczas krystalizacji przesiewowych wstępne warunki krystalizacyjne trzykrotnie optymalizowałam, wykorzystując cztery różne preparaty białka. Pierwszą optymalizacje prowadziłam dla warunków fosforanowych i preparatu B (15 mg/mL białko AnsA, 20 mM Tris pH 7.5, 5 mM β-merkaptoetanol). Ponowna optymalizacja, która była prowadzona jednocześnie dla dwóch preparatów zawierających Tris pH 7.5 i TCEP, ale różniących się obecnością lub brakiem L-Asp, dotyczyła już trzech typów warunków (fosforanowe, siarczanowe, octanowe). Trzecie podejście do optymalizacji trzech typów warunków wynikało z pomysłu, że zmiana pH preparatu białka (który zawierał $(K_2H/NaH_2)PO_4$ pH 9.0 i TCEP) może wpłynąć na jakość dyfrakcji, co jednak nie sprawdziło się.

Spośród ośmiu warunków krystalizacyjnych (podzielonych przeze mnie na trzy

typy: fosforanowe, siarczanowe, octanowe) tylko jeden udało mi się zoptymalizować na tyle, by móc przeprowadzić rejestrację danych dyfrakcyjnych. Był to warunek octanowy, konkretnie nr 12 z zestawu Clear Strategy Screen II (MD1-15) firmy MDL, który w połączeniu z oktyloglukozą, pozwalał uzyskiwać kryształy w postaci rozdzielonych płytek. Co ciekawe znajdujące się w tym warunku jony Ca^{2+}, ewidentnie poprawiające morfologię kryształów, znane są z tego, że powodują obniżenie aktywności asparaginazy z *Rhizobium etli* (**Moreno-Enriquez *et al.*, 2012**). Najlepsze uzyskiwane przeze mnie kryształy białka AnsA rozpraszały promieniowanie synchrotronowe do rozdzielczości 2.2 Å i tylko raz udało mi się uzyskać kryształ rozpraszający do 1.5 Å.

6.3.4. Cechy kryształów a jakość dyfrakcji

Kryształy, jakie wyrosły w zmodyfikowanym warunku octanowym, opracowanym dzięki zakrojonej na szeroką skale optymalizacji, były pokroju płytkowego, stosunkowo duże i silnie rozpraszały promieniowanie rentgenowskie. Jednocześnie, jak wynika z przeprowadzonej weryfikacji, jakość zarejestrowanych danych zmieniała się w zależności od kierunku padania wiązki promieniowania na kryształ. Mimo tej znaczącej wady, wszystkie zarejestrowane w niniejszym projekcie dane pochodzą zasadniczo tylko z tego jednego rodzaju kryształów, ponieważ były one jedynymi, które nadawały się do eksperymentów dyfrakcyjnych i lepszych, mimo wielu prób, nie uzyskałam.

W związku z przyjętą strategią, która zakładała wykorzystanie sygnału anomalnego do rozwiązania struktury krystalicznej, kryształy natywnego białka AnsA nasączałam kompleksem tantalu $[Ta_6Br_{12}]Br_2$ celem wprowadzenia do nich atomów specjalnych. Niestety taka ingerencja w skład kryształów znacząco pogarszała ich zdolność rozpraszania, jeszcze bardziej komplikując sytuację. Wynik był taki, że przeprocesowane dane nie zawierały satysfakcjonującego sygnału anomalnego pochodzącego od atomów tantalu. Zaniechałam zatem dalszych prób nasączania kryształów (również innymi związkami) i postanowiłam podjąć próbę zarejestrowania sygnału anomalnego pochodzącego od naturalnie występujących w białku atomów siarki. Korzystając z natywnych kryształów AnsA, zarejestrowano kilka zbiorów danych, naświetlając kryształy promieniowaniem o długości fali λ ok. 2 Å. Niestety kryształy te źle znosiły długą i wielokrotną ekspozycje na promieniowanie rentgenowskie, której celem było uzyskanie dużej redundancji, niezbędnej do

uchwycenia naturalnie słabego sygnału pochodzącego od atomów siarki. I w tych przypadkach anizotropia kryształów uniemożliwiła uzyskanie odpowiednich danych.

Trzeci rodzaj atomów specjalnych, jakie planowałam wykorzystać do uzyskania sygnału anomalnego, to atomy selenu. Wprowadzałam je do struktury białka AnsA już na etapie jego produkcji. Mimo, że właściwości preparatu AnsA-Se (rozpuszczalność, stabilność, widmo DLS) były podobne do tych zaobserwowanych dla natywnej wersji enzymu, białko to stanowiło pochodną, która potencjalnie mogła krystalizować zupełnie inaczej, co z kolei dawało nadzieje na kryształy lepszej jakości. Niestety oczekiwana zmiana nie przyniosła pozytywnych efektów, ponieważ elastyczne kryształy AnsA-Se w ogóle nie rozpraszały promieniowania rentgenowskiego.

Spośród wszystkich zarejestrowanych zbiorów danych dyfrakcyjnych do dalszych prac nadaje się tylko jeden, uzyskany podczas naświetlania kryształu natywnego białka AnsA standardowym promieniowaniem o długości 0.91841 Å. Użyty kryształ również rozpraszał anizotropowo, aczkolwiek jednocześnie do najwyższej rozdzielczości (~1.5 Å). Taki zestaw danych może posłużyć w przyszłości do rozwiązania struktury metodą MR. Co prawda brak podobnej struktury, która mogłaby zostać zastosowana jako model, znacznie utrudnia zadanie. Jednak bardzo szybki rozwój narzędzi bioinformatycznych daje nadzieję na wygenerowanie modelu teoretycznego, który będzie wystarczająco dobry do podstawienia cząsteczkowego.

Moim zadaniem w projekcie dotyczącym asparaginazy z *Rhizobium etli* było przeprowadzenie obiektu badań przez wszystkie etapy dotyczące pracy z preparatem białkowym, umożliwiające podjęcie próby rozwiązania pierwszej struktury tego typu. Kończąc swoje zadanie przekazałam wszystkie zdobyte informacje i dane osobie, która aktualnie pracuje nad rozwiązaniem struktury.

7. Potasowo-zależna asparaginaza z *Phaseolus vulgaris*

7.1. Jeden enzym w trzech wersjach

Prace nad asparaginazą potasowo-zależną (PvAspG1) mają swój początek w Kanadzie, gdzie grupa dr Frederica Marsolais stworzyła plazmidowy wektor ekspresyjny (pProExHTb) zawierający cDNA o sekwencji transkryptu Phvul.001g025000.1 z dojrzewających nasion fasoli zwyczajnej (*Phaseolus vulgaris*). Wyprodukowane i oczyszczone białko zostało także poddane badaniom kinetycznym, które potwierdziły zależność aktywności asparaginazowej enzymu od obecności jonów potasu (**Bejger *et al*., 2014**).

Wektor ekspresyjny (pProExHTb) został następnie przekazany prof. Mariuszowi Jaskólskiemu, kierownikowi Centrum Badań Biokrystalograficznych IChB PAN w Poznaniu. W CBB dr Barbara Imiołczyk wyprodukowała, oczyściła i wykrystalizowała białko. Podczas przygotowania preparatu białka do krystalizacji analiza elektroforetyczna SDS-PAGE potwierdziła, że PvAspG1 ulega procesowi dojrzewania, jaki zachodzi również w przypadku roślinnych asparaginaz potasowo-niezależnych, a który polega na autokatalitycznym cięciu łańcucha polipeptydowego na podjednostki α i β. Stwierdzono również, że obie podjednostki mają spodziewaną masę 21 kDa (α) i 13.6 kDa (β).

Aby ustalić molekularne/strukturalne podłoże zależności tej asparaginazy od jonów potasu, stworzone zostały trzy wersje preparatu białka różniące się zawartością jonów metali: jedna przy użyciu samych jonów potasu (PvAspG1-K), jedna przy użyciu samych jonów sodu (PvAspG1-Na) oraz jedna przy użyciu obu rodzajów jonów (PvAspG1-K/Na) **(Bejger *et al*., 2014**). Wszystkie trzy preparaty wykrystalizowały w roztworze zawierającym azotan sodu (Tabela 7.1). W przypadku preparatu PvAspG1-K, chcąc uniknąć zmieszania jonów K^+ i Na^+ w kropli krystalizacyjnej, podjęto również próby wykrystalizowania białka w azotanie potasu, który jednak nie indukował nukleacji. Najlepszym z przetestowanych krioprotektantów okazał się również $NaNO_3$. Ponownie, aby nie dopuścić do niepotrzebnego dodania jonów sodu, pomiary dyfrakcji kryształu PvAspG1-K zostały wykonane inaczej niż PvAspG1-Na i PvAspG1-K/Na. W przypadku PvAspG1-K naświetlano kryształ „osuszony", ponieważ przetestowane krioprotektanty (np. glicerol, PEG400, itp.) dawały gorsze efekty. Dane dyfrakcyjne dla kryształów PvAspG1-K, PvAspG1-Na i PvAspG1-K/Na zostały zarejestrowane na

wiązce BL14.2 synchrotronu BESSY w Berlinie. Szczegółowe informacje na temat przebiegu pomiarów dyfrakcji zebrałam w Tabeli 7.2.

Tabela. 7.1. Protokół oczyszczania i krystalizacji trzech wersji białka PvAspG1 (pogrubioną czcionką zaznaczyłam zastosowane jony metali alkaicznych)

	PvAspG1-K	PvAspG1-K/Na	PvAspG1-Na
odczynniki używane podczas oczyszczania	50 mM Tris pH 8.0 20 - 300 mM imidazol 150 - 500 mM **K**Cl		50 mM Tris pH 8.0 20 - 300 mM imidazol 150 - 500 mM **Na**Cl
skład preparatu białka do krystalizacji	stężenie białka: 20 mg/mL 50 mM Tris pH 8.0 1 mM TCEP 150 mM **K**Cl		stężenie białka: 20 mg/mL 50 mM Tris pH 8.0 1 mM TCEP 150 mM **Na**Cl
dodany do preparatu ligand - sól kwasu asparaginowego (nadmiar molowy)	**K**-Asp (1:3)	**Na**-Asp (1:25)	**Na**-Asp (1:25)
skład roztworu krystalizacyjnego	100 mM bis-tris propan pH 8.5 20% (w/v) PEG3350 200 mM **Na**NO_3		
zastosowana krioprotekcja	-	2 M **Na**NO_3	2 M **Na**NO_3

Tabela 7.2. Informacje dotyczące pomiarów dyfrakcji trzech kryształów PvAspG1.

	PvAspG1-K	PvAspG1-K/Na	PvAspG1-Na
źródło promieniowania	BL14.2 BESSY Berlin	BL14.2 BESSY Berlin	BL14.2 BESSY Berlin
długość fali (Å)	0.918	0.827	0.918
temperatura pomiarów (K)	100	100	100
oscylacja (º)	0.5	1.0	0.5
liczba obrazów	370	100	540

7.2. Część doświadczalna

7.2.1. Rozwiązywanie, udokładnianie i ocena jakości struktur PvAspG1

Ze względu na duże podobieństwo sekwencji aminokwasowych asparaginaz typu roślinnego (ok. 70%) oraz obecność struktur asparaginaz potasowo-niezależnych w bazie PDB, w przypadku badanego przeze mnie enzymu potasowo-zależnego z fasoli zwyczajnej do rozwiązania problemu fazowego mogłam zastosować metodę podstawienia cząsteczkowego. Jako pierwszą rozwiązałam i udokładniłam strukturę PvAspG1-K, następnie PvAspG1-Na. Struktura PvAspG1-K/Na została rozwiązana i wstępnie udokładniona przez Damiena Clavel, studenta stażystę w CBB. Końcowe

udokładnianie, walidacja i deponowanie modeli w banku PDB, jak również ich szczegółowa analiza porównawcza zostały przeprowadzone przeze mnie.

7.2.1.1. Strategia, modele, użyte programy

Zarejestrowany dla PvAspG1-K zbiór danych dyfrakcyjnych został przeze mnie wywskaźnikowany i zintegrowany za pomocą programu *XDS* (**Kabsch, 2010**). W trakcie procesowania ustalałam limit rozdzielczości tak, by stosunek sygnału do szumu <I/σI> w ostatnim przedziale rozdzielczości był bliski 2. Symetria sieci i wygaszenia systematyczne wskazały grupę przestrzenną $P2_12_12_1$. Szczegółowe informacje uzyskane podczas procesowania i skalowania danych zebrałam w Tabeli 7.3. Otrzymany w postaci pliku *hkl* zbiór wskaźników refleksów i ich intensywności wraz z wartościami odchyleń standardowych przekształciłam w zbiór amplitud czynników struktury w formacie *mtz* przy pomocy programu *XDSCONV* (**Kabsch, 2010**). Na tym etapie wyodrębniłam także podzbiór refleksów testowych wyłączonych z udokładniania (Tabela 7.3), na podstawie których w trakcie udokładniania obliczałam wskaźnik rozbieżności R_{free} (**Brünger, 1992**).

Strukturę kryształu PvAspG1-K rozwiązałam stosując program *PHASER* (**McCoy *et al*., 2007**). Do tego zadania niezbędna była informacja o zawartości niezależnej części komórki elementarnej (ASU), którą uzyskałam na podstawie analizy współczynnika Matthewsa (**Matthews, 1968**). Analiza ta sugerowała, że w części niezależnej ASU obecne są dwie cząsteczki typu (αβ), (V_M = 2.68 $Å^3$/Da, Tabela 7.3). Do podstawienia cząsteczkowego użyłam modelu asparaginazy z łubinu żółtego (*Lupinus luteus*, LlA) zdeponowanego w bazie PDB z kodem PDB 2gez. Poprawność rozwiązania potwierdziły wyniki obliczeń dwóch parametrów: (I) przyrost logarytmu wiarogodnosci (*ang. log-likelihood-gain*) LLG = 2057 oraz powiązany z nim (II) z-score = 43.8. Po wstępnym udokładnianiu w programie *PHENIX* (**Adams *et al*., 2010**), wskaźnik rozbieżności (R) wynosił 30%.

Udokładnianie struktury PvAspG1-K prowadziłam w programie *PHENIX* (**Adams *et al*., 2010)** z zastosowaniem parametrów *TLS* wyznaczonych przez program *TLSMD* (**Painter i Merrit, 2006**). Procedura obejmowała ręczną przebudowę modelu w oparciu o mapy gęstości elektronowej ($2F_o$-F_c i F_o-F_c) w programie *COOT* (**Emsley i Cowtan, 2004**). Początkowo model PvAspG1-K posiadał błędną sekwencję pochodzącą z modelu użytego do podstawienia cząsteczkowego, dlatego przy użyciu programu

ARP/wARP (**Langer *et al.*, 2008**) zmieniałam reszty aminokwasowe zgodnie z sekwencją białka PvAspG1 z *Phaseolus vulgaris* zdeponowaną w bazie PHYTOZOME (www.phytozome.net). Szczegółowe informacje na temat przebiegu procesu udokładniania struktury oraz statystyki ostatecznych wyników zebrałam w Tabeli 7.3.

W przypadku PvAspG1-Na podczas procesowania danych dyfrakcyjnych postępowałam analogicznie jak poprzednio i używałam tych samych programów (Tabela 7.3). Strukturę rozwiązałam również w programie *PHASER*, jednak do podstawienia cząsteczkowego wykorzystałam model PvApsG1-K. Wskaźnik rozbieżności (R) rozwiązanej struktury PvAspG1-Na wynosił 24%. Tym razem model asparaginazy udokładniałam w programie *REFMAC5* (**Murshudov *et al.*, 2011**) z platformy *CCP4i* (**Winn *et al.*, 2011**). Ponieważ jako sondę do rozwiązania struktury metodą podstawienia cząsteczkowego zastosowałam to samo białko, przebudowa modelu w programie *COOT* nie wymagała „mutowania" reszt aminokwasowych i polegała przede wszystkim na zmianie konformacji łańcuchów bocznych. W tym przypadku udokładnianie prowadziłam z wykorzystaniem z parametrów TLS, wyznaczonych przez program *REFMAC5* standardowo definiujący jedną grupę TLS dla jednego łańcucha polipeptydowego (Tabela 7.3). Kontynuując udokładnianie trzeciego modelu PvAspG1-K/Na korzystałam z programu *PHENIX* i *COOT* oraz serwera *TLSMD* (Tabela 7.3).

Aby mieć pewność, że trzy udokładniane modele potasowo-zależnej asparaginazy z fasoli zwyczajnej (PvAspG1-K, PvAspG1-Na, PvAspG1-K/Na) są zgodne z danymi eksperymentalnymi, wyniki kolejnych cykli obliczeń weryfikowałam opierając się na dostępnej literaturze (**Wlodawer *et al.*, 2008**). Podstawowymi sprawdzanymi parametrami były wskaźniki rozbieżności: R (którego wartość starałam się maksymalnie obniżyć) oraz R_{free} (którego wzrost ostrzegał mnie o nadinterpretacji danych i map gęstości elektronowej). Co więcej, analizowałam średnie kwadratowe odchylenia od wartości standardowych dla długości wiązań (r.m.s.d. B) i kątów walencyjnych (r.m.s.d. A), jak również wartości czynników temperaturowych poszczególnych atomów (B) i rozkład kątów torsyjnych łańcucha głównego na wykresie Ramachandrana. Podczas weryfikacji powyższych parametrów korzystałam z programów Procheck (**Laskowski *et al.*, 1993**) i MolProbity (**Chen *et al.*, 2010**) dostępnych na serwerze ADIT.

Tabela 7.3. Charakterystyka trzech zbiorów danych krystalograficznych oraz trzech udokładnianych modeli struktury PvAspG1. W nawiasach podano wartości dotyczące ostatniego zakresu rozdzielczości.

	PvAspG1-K	PvAspG1-K/Na	PvAspG1-Na
kod PDB	4pu6	4pv2	4pv3
dane dyfrakcyjne			
grupa przestrzenna	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
parametry komórki elementarnej (Å)	a 56.94 b 102.67 c 127.04	a 56.88 b 102.63 c 123.75	a 57.40 b 103.42 c 124.56
mozaikowość kryształu (º)	0.24	0.21	0.38
zakres rozdzielczości (Å)	47.60 – 2.30 (2.43 – 2.30)	38.77 – 1.79 (1.89 – 1.79)	46.55 – 2.09 (2.21 – 2.09)
liczba refleksów zarejestrowanych	239 668 (27 620)	280 364 (44 366)	481 561 (76 242)
liczba refleksów niezależnych	32 518 (4 237)	68 445 (10 836)	44 381 (7 009)
kompletność (%)	96.3 (78.8)	99.1 (98.2)	99.6 (99.0)
krotność pomiaru	7.37 (6.52)	4.10 (4.09)	10.85 (10.88)
$\langle I/\sigma(I)\rangle$	22.97 (1.98)	16.33 (1.98)	15.78 (2.47)
R_{merge}	0.067 (0.903)	0.053 (0.630)	0.104 (0.973)
udokładnianie			
użyty program	*PHENIX*	*PHENIX*	*REFMAC5*
zawartość ASU	$(\alpha\beta)_2$	$(\alpha\beta)_2$	$(\alpha\beta)_2$
współczynnik Matthewsa ($Å^3$/Da)	2.68	2.61	2.67
liczba grup TLS [łańcuch A / B / C / D]	22 [4/7/6/5]	29 [9/8/5/7]	4 [1/1/1/1]
liczba refleksów w zbiorze roboczym / testowym	31 509 / 1 009	67 213 / 1 232	43 451 / 930
R / R_{free} (%)	17.59 / 23.13	16.35 / 20.66	17.80 / 23.14
liczba elementów / średni czynnik B ($Å^2$) - atomy białka - cząsteczki wody - jony K^+ - jony Na^+ - L-Asp	 4 234 / 54.2 114 / 42.4 4 / 48.0 0 / - 1 / 39.5	 4 286 / 30.6 421 / 37.4 4 / 36.3 4 / 21.3 0 / -	 4 240 / 44.9 286 / 47.2 0 / - 4 / 42.3 0 / -
r.m.s.d. dla - długości wiązań (Å) - kątów walencyjnych (º)	 0.018 1.47	 0.019 1.37	 0.019 1.79
statystyki wykresu Ramachandrana (%) - konformacja uprzywilejowana - konformacja dozwolona	 94.89 5.11	 97.18 2.82	 97.00 3.00

7.2.1.2. Niebiałkowe składniki w strukturach kryształów

W trakcie udokładniania, model PvAspG1-K uzupełniałam o cząsteczki wody, opierając się na kształcie różnicowych map gęstości elektronowej i regule mówiącej, że dobrze wymodelowana cząsteczka wody powinna mieć maksymalnie cztery (a

przynajmniej jedno) tetraedrycznie rozmieszczone wiązania wodorowe z molekułą białka lub z inną „dobrą" cząsteczką wody. Ostatecznie zidentyfikowałam 114 cząsteczek H_2O. Jednocześnie, na podstawie wyjątkowo silnego sygnału na mapie gęstości elektronowej i dużej liczby kontaktów międzyatomowych zlokalizowałam 4 jony potasu, w pozycjach oznaczonych jako: S_A, S_C, A_A, A_C (znaczenie tych nazw wyjaśniam później). Ponadto, w oparciu o dodatnią gęstość różnicową, w modelu PvAspG1-K w miejscu aktywnym zidentyfikowałam jedną cząsteczkę liganda, którym jest produkt reakcji - L-asparaginian, dodawany do preparatu białka przed krystalizacją. Po kolejnych cyklach udokładniania, gdy jakość map gęstości elektronowej ulegała znacznej poprawie, a w okolicy liganda pojawiła się (znikoma) ujemna gęstość różnicowa. Wbudowanie kilku cząsteczek wody zamiast L-Asp nie było możliwe, ze względu na zbyt małe odległości pomiędzy pikami, dlatego zmniejszyłam obsadzenie liganda L-Asp do 70%, co pozwoliło w pełni zinterpretować istniejącą gęstość elektronową i nie powodowało pojawiania się ujemnej gęstości różnicowej.

Proces udokładniania kolejnych dwóch modeli asparaginazy przebiegł w podobny sposób. W przypadku PvAspG1-Na zidentyfikowałam 286 cząsteczek wody. W pozycjach analogicznych do pozycji jonów potasu modelu PvAspG1-K (S_A, S_C, A_A, A_C) wymodelowałam 4 jony sodu. Nie zaobserwowałam jednak wystarczająco klarownej różnicowej gęstości, która mogłaby świadczyć o obecności L-asparaginianu w centrum aktywnym, mimo obecności asparaginianu sodu w 25-krotnym nadmiarze podczas krystalizacji (Tabela 7.1). W przypadku PvAspG1-K/Na dzięki znacznie wyższej rozdzielczości, mogłam wymodelować aż 421 cząsteczek wody i dodatkowo 21 jonów azotanowych, dodanych do modelu na podstawie kształtu gęstości elektronowej i informacji o obecności tych jonów w roztworze krystalizacyjnym i w krioprotektancie. Model uzupełniłam czterema jonami potasu i czterema jonami sodu dzielącymi te same pozycje (S_A, S_C, A_A, A_C) w proporcjach 30% i 70%. Również w tym przypadku nie znalazłam cząsteczki L-Asp w centrum aktywnym.

7.2.1.3. Identyfikacja jonów metali w strukturach

W celu potwierdzenia obecności i rodzajów jonów metali w wyznaczonych pozycjach w strukturach PvAspG1 (S_A, S_C, A_A, A_C) przeprowadziłam analizę dowodów bezpośrednich i pośrednich.

Posłużyłam się między innymi bardzo słabym sygnałem anomalnym obecnym w danych natywnych zarejestrowanych dla PvAspG1. Przy długościach fal, które wykorzystano (0.918 i 0.827 Å) poprawki anomalne *f"* czynników rozpraszania potasu i sodu przyjmują zbyt małe wartości (*f"*(K) ≈ 0.4 e, *f"*(Na) ≈ 0.04 e) aby można było uzyskać sygnał przydatny do rozwiązania problemu fazowego. Jednak w przypadku potasu efekt ten może być wystarczający do lokalizacji atomów rozpraszających anomalnie. Za pomocą programu *XDS* (**Kabsch, 2010**) ponownie przeprocesowałam dane dyfrakcyjne bez uśredniania par Bijvoeta, po czym za pomocą programu *SHELXC* (**Sheldrick, 2008**) uzyskałam czynniki struktury. Następnie wykorzystując program *Anode* **(Thorn i Sheldrick, 2011)** policzyłam anomalne mapy gęstości elektronowej i zweryfikowałam je nakładając na modele białka w programie COOT (**Emsley i Cowtan, 2004)**. W przypadku PvAspG1-Na i -K/Na słabe, dające się zinterpretować piki zauważyłam jedynie przy atomach siarki. Wnioskuję zatem, że w tych kryształach nie ma (przynajmniej w pełnym obsadzeniu) atomów pierwiastków cięższych niż siarka. Przeprowadzona przeze mnie podobna analiza dla PvAspG1-K wykazała obecność pików anomalnych w pobliżu czterech jonów potasu, o sile nieco większej (2 - 4 σ) niż piki przy atomach siarki (~ 2 σ). Eksperyment ten pozwala stwierdzić, że zidentyfikowane gęstości elektronowe należą do atomów o większej niż siarka liczbie elektronów, dlatego obecność jonów potasu w tych miejscach jest znacznie bardziej prawdopodobna niż jonów sodu.

Aby znaleźć kryteria pozwalające odróżnić od siebie jony potasu i sodu w strukturach białek przeanalizowałam dostępną literaturę. Zarówno długość wiązań ligand (O) - metal, ich liczba, jak i geometria sfery koordynacyjnej zależą przede wszystkim od rodzaju metalu. Należy jednak pamiętać, że im więcej wiązań koordynacyjnych tworzy jon centralny tym są one dłuższe. Ważny jest także rodzaj atomu liganda, jako że odległość jon metalu – tlen cząsteczki wody zazwyczaj jest dłuższa niż odległość jon metalu – tlen karbonylowy łańcucha głównego białka (**Harding *et al.*, 2010**). Rozdzielczość jest kolejnym istotnym czynnikiem, ponieważ im jest niższa tym większe są rozbieżności pomiędzy występującymi w modelu długościami wiązań, a typowymi dla danego jonu wartościami standardowymi. To samo dotyczy kątów walencyjnych.

Oba jony, potasowy i sodowy, częściej koordynowane są przez atomy tlenu grupy karbonylowej łańcucha głównego, niż przez łańcuchy boczne reszt aminokwasowych i dla obu jonów liczba koordynacyjna najczęściej wynosi pięć lub sześć (**Harding *et al.*,**

2010). Typowe długości wiązań koordynacyjnych w tego typu oddziaływaniach zostały ustalone (**Harding, 2006; Harding *et al.*, 2010; Kuppuraj *et al.*, 2009**) na podstawie analizy struktur zdeponowanych w bazach danych PDB (**Berman *et al.*, 2000**) i CSD (**Allen, 2002**). Oszacowane odległości pomiędzy jonem metalu i tlenem łańcucha głównego białka wynoszą 2.38 (10) Å dla Na^+ i 2.74 (15) Å dla K^+, a pomiędzy jonem metalu i tlenem cząsteczki wody 2.41 (10) Å dla Na^+ i 2.81 (15) Å dla K^+ (**Harding *et al.*, 2010**).

Po wymodelowaniu i udokładnieniu poprawność obsadzenia jonu metalu weryfikowałam przeprowadzając test BVM (Bond Valence Method, **Brese i O'Keeffe, 1991**) oraz korzystając z serwera CMM (CheckMyMetal, **Zheng *et al.*, 2014**). Oczywiście, wiarygodność wyników takich obliczeń jest tym większa im wyższa jest rozdzielczość, ponieważ analizowane parametry są wtedy bardziej dokładne. Według niektórych opinii niepodważalne wyniki można uzyskać analizując struktury o rozdzielczości ok. 1.5 Å i wyższej (**Brown, 2009; Muller, 2003**).

Podczas przygotowywania preparatu białka PvAspG1-K używano jedynie soli potasu, natomiast roztwór krystalizacyjny zawierał jony sodu (w postaci $NaNO_3$), dlatego teoretycznie jony obu metali mogłyby być obecne w strukturze. W modelu PvAspG1-K cztery pozycje metali obsadziłam jednak jonami potasu, ze względu na wskazanie map anomalnych oraz silny sygnał pików gęstości elektronowej $2F_o$-F_c w tych miejscach (4.1 - 6.3 σ), sugerujący obecność elementów cięższych niż jony sodu czy cząsteczki wody. Zamiana jonów potasu na jony sodu skutkowała pojawieniem się dodatniej różnicowej gęstości elektronowej w okolicach trzech pozycji metali: S_A, A_A, A_C. Wykonany test BVM wskazuje, że wszystkie jony K^+ wymodelowałam poprawnie (Tabela 7.4). Program CMM również potwierdza obecność jonów potasu (Rysunek 7.1). Biorąc pod uwagę wszystkie dowody bezpośrednie i pośrednie można z bardzo dużym prawdopodobieństwem wnioskować, o obecności jonów K^+ we wszystkich czterech pozycjach modelu PvAspG1-K.

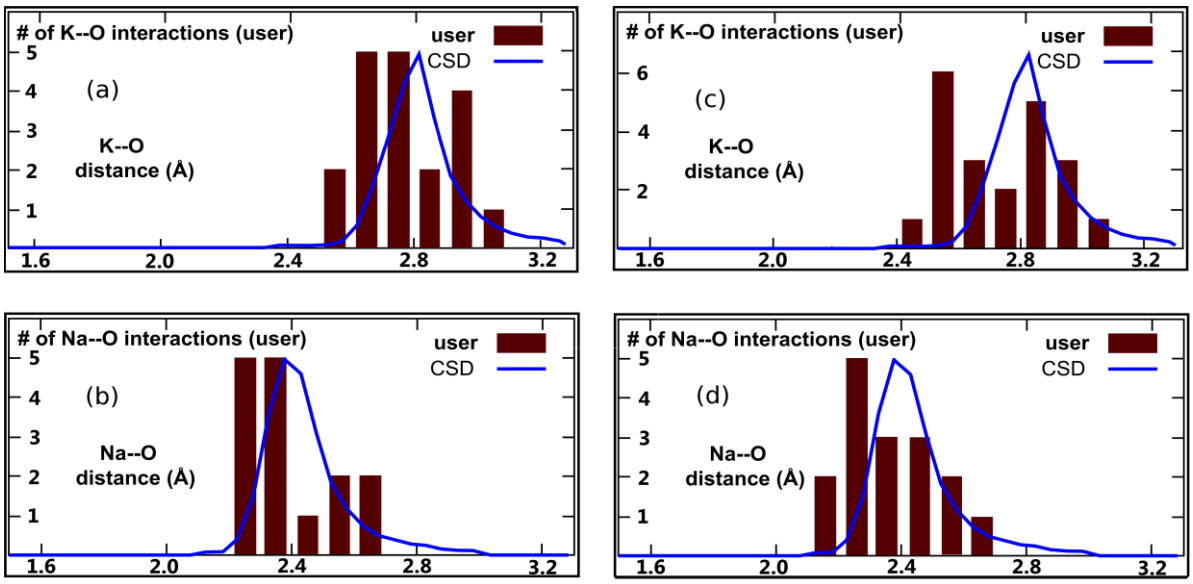
W przypadku PvAspG1-Na przygotowywanie preparatu białka, krystalizacja i pomiary dyfrakcji odbywały się w obecności wyłącznie jednego rodzaju jonów, jonów sodu, dlatego prawdopodobieństwo obecności innego kationu w tej strukturze jest bardzo nikłe. Piki mapy gęstości elektronowej $2F_o$-F_c w czterech pozycjach jonów metali w modelu są na poziomie sygnału odpowiadającego cząsteczkom wody (2.9 – 4.0 σ). Wymodelowane cztery jony sodu pasują we wspominanych pozycjach ze względu na liczbę elektronów oraz charakter oddziaływań z atomami ligandów (sześć i

więcej oddziaływań o długości znacznie krótszej niż typowa długość wiązań wodorowych O···H_2O). Zamiana jonów sodu na potas nie skutkowała co prawda pojawieniem się ujemnej gęstości różnicowej w okolicach pozycji metali, jednak test BVM oraz analiza CMM wskazują, że we wszystkich czterech pozycjach metalem jest jon Na^+ (Tabela 7.4., Rysunek 7.1). W świetle przeprowadzonych analiz obecność jonów sodu we wszystkich czterech pozycjach modelu PvAspG1-Na nie budzi wątpliwości.

Tabela 7.4. Wyniki testu BVM dla wymodelowanych jonów metali trzech struktur PvAspG1 (wyniki dla czterech pozycji podałam w kolejności: górny wiersz: S_A / A_A, dolny wiersz: S_C / A_C). Zacieniowane komórki tabeli podają wyniki otrzymane przy założeniu obecności w danej pozycji alternatywnego jonu (K^+ zamiast Na^+ i odwrotnie). Poprawny parametr R_{ij} powinien być zgodny z wartością literaturową zawartą w Tabeli 5. Brese i O'Keeffe (1991) i wynosić odpowiednio: R_{KO}(K) = 2.13, R_{NaO}(Na) = 1.80. V_i podaje obliczoną walencyjność metali (dla obu jonów powinna wynosić oczywiście ok. 1.0).

	koordynowany metal	R_{ij}	V_i metalu przy założeniu obecności:	
			K^+	Na^+
PvAspG1-K	K^+	2.01 / 2.09 2.08 / 2.01	1.11 / 1.21 1.39 / 1.14	0.45 / 0.50 0.57 / 0.47
PvAspG1-Na	Na^+	1.80 / 1.78 1.78 / 1.92	2.43 / 2.80 2.55 / 1.76	0.99 / 1.06 1.05 / 0.72
PvAspG1-K/Na	K^+	2.08 / 2.09 2.08 / 2.34	1.15 / 1.10 1.15 / 0.56	0.47 / 0.45 0.47 / 0.23
	Na^+	1.76 / 1.82 1.74 / 1.87	2.71 / 2.32 2.86 / 2.04	1.10 / 0.95 1.18 / 0.94

Rysunek 7.1. Wyniki analizy czterech pozycji jonów metali (S_A, S_C, A_A, A_C) modeli: (a) PvAspG1-K, (b) PvAspG1-Na i (c,d) PvAspG1-K/Na przeprowadzonej przeze mnie w programie CMM.



Sytuacja jest bardziej skomplikowana w przypadku modelu PvAspG1-K/Na, ponieważ w tej strukturze w pozycjach metali najprawdopodobniej występują oba jony

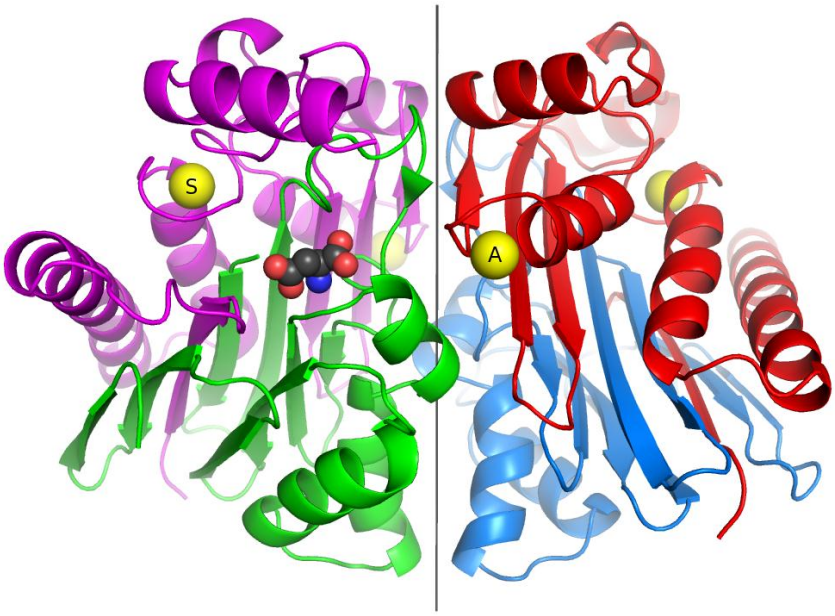
jednocześnie. Jest to możliwe, ponieważ podczas oczyszczania, krystalizacji i pomiarów dyfrakcji używano obu jonów. W tym przypadku, pomimo najlepszej rozdzielczości (1.79 Å) niejednoznaczne mapy gęstości elektronowej są stosunkowo trudne do interpretacji. Starając się jak najtrafniej zinterpretować gęstość elektronową w każdej z czterech pozycji metalu (S_A, S_C, A_A, A_C) wymodelowałam oba jony. Następnie metodą prób i błędów, obserwując czynniki temperaturowe oraz mapy $2F_o$-F_c i F_o-F_c, ustaliłam cząstkowe obsadzenia metali jako 30% potas i 70% sód. Postępując tak uzyskałam model najbardziej przybliżony do rzeczywistości. Test BVM (Tabela 7.4) i analiza CMM (Rysunek 7.1) potwierdziły taką interpretację. Co więcej, podwójne obsadzenie jonami K^+ i Na^+ koreluje z wynikami analizy strukturalnej poszczególnych modeli tej asparaginazy.

7.2.2. Analiza trzech modeli PvAspG1

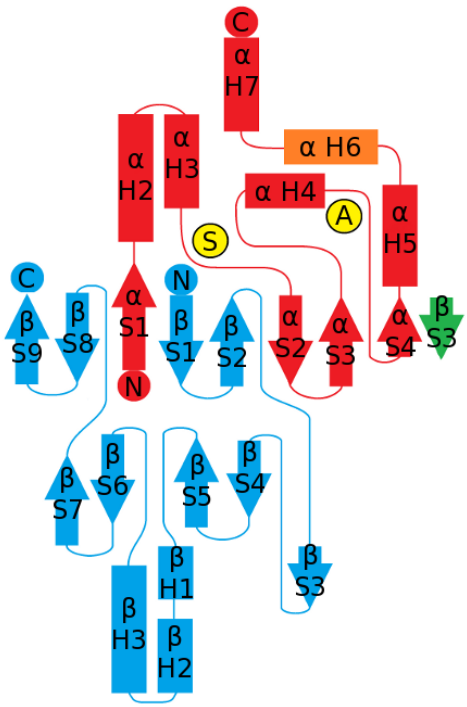
Przeprowadzając analizę struktury asparaginazy potasowo-zależnej z fasoli zwyczajnej, którą wykonywałam na podstawie wszystkich trzech przygotowanych modeli, korzystałam z wielu programów, takich jak np. BLAST (**Altschul *et al.*, 1997**), DSSP (**Kabsh i Sander, 1983**), ALIGN (**Satow *et al.*, 1986**), PDBsum (**Laskowski, 2009**) czy PyMOL (**DeLano, 2002**). Jednak najwięcej cennych wniosków wysnułam już podczas modelowania w programie COOT, w którym wykorzystując narzędzie SSM (**Krissinel i Henrick, 2004**) porównywałam różne modele PvAspG1.

7.2.2.1. Ogólna charakterystyka struktury PvAspG1

Podobnie jak inne asparaginazy typu roślinnego, enzym potasowo-zależny z fasoli zwyczajnej powstaje z dwóch identycznych łańcuchów polipeptydowych. Oba łańcuchy podlegają procesowi dojrzewania, który polega na autoproteolitycznym przecięciu na podjednostki α i β. Dojrzałe białko jest zatem heterotetramerem zbudowanym z czterech łańcuchów białkowych (A, B, C, D) zwiniętych w dwa heterodimery αβ (AB i CD) związane niekrystalograficzną osią dwukrotną (Rysunek 7.2). Heterodimery αβ PvAspG1 posiadają typ pofałdowania bardzo podobny do asparaginaz potasowo-niezależnych. Tworzą one tzw. "kanapkę αββα", czyli strukturę, której rdzeń stanowią dwa arkusze β, otoczone z obu stron helisami α.



Rysunek 7.2. Struktura czwartorzędowa PvAspG1. Rysunek wykonałam na podstawie PvAspG1-K. Łańcuchy polipeptydowe oznaczyłam kolorami: A - fioletowy (podjednostka α), B - zielony (β), C - czerwony (α), D - niebieski (β); L-Asp w kolorze czarno (C) – czerwono (O) – niebieskim (N) ustala pozycję miejsca aktywnego; żółte kule reprezentują jony metalu w pozycjach A i S; pionowa linia oznacza niekrystalograficzną oś dwukrotną.



Rysunek 7.3. Diagram topologiczny heterodimeru αβ struktury PvAspG1. Podjednostkę α (łańcuch A lub C) oznaczyłam kolorem czerwonym, podjednostkę β (B lub D) niebieskim. Pomarańczowy element podjednostki α oznacza helisę 3_{10}, zielony element, będący ostatnią nicią arkusza β, pochodzi z sąsiedniego heterodimeru αβ. Żółte kule reprezentują jony metalu w pozycjach $A_{A/C}$ i $S_{A/C}$. Trzyczęściowe etykiety elementów struktury drugorzędowej oznaczają: α lub β – podjednostki heterodimeru, H – helisę, S - łańcuch β, numer elementu zgodny z topologią asparaginazy z łubinu żółtego LlA (**Michalska *et al.*, 2006**).

Diagram topologiczny (Rysunek 7.3), który przygotowałam opierając się na diagramie struktury asparaginazy z łubinu żółtego (LlA) i wyniku analizy struktury PvAspG1 w programie DSSP (**Kabsh i Sander, 1983**), uwidacznia ułożenie wszystkich elementów struktury drugorzędowej. Kolejnym elementom struktury drugorzędowej nadałam nazwy dla lepszego ich rozróżnienia. Nazwy precyzują, gdzie znajduje się element (w podjednostce α lub β) oraz jakiego jest rodzaju (H - helisa α, S - łańcuch β), natomiast numeracja elementów pozostała zgodna z numeracją diagramu topologicznego LlA.

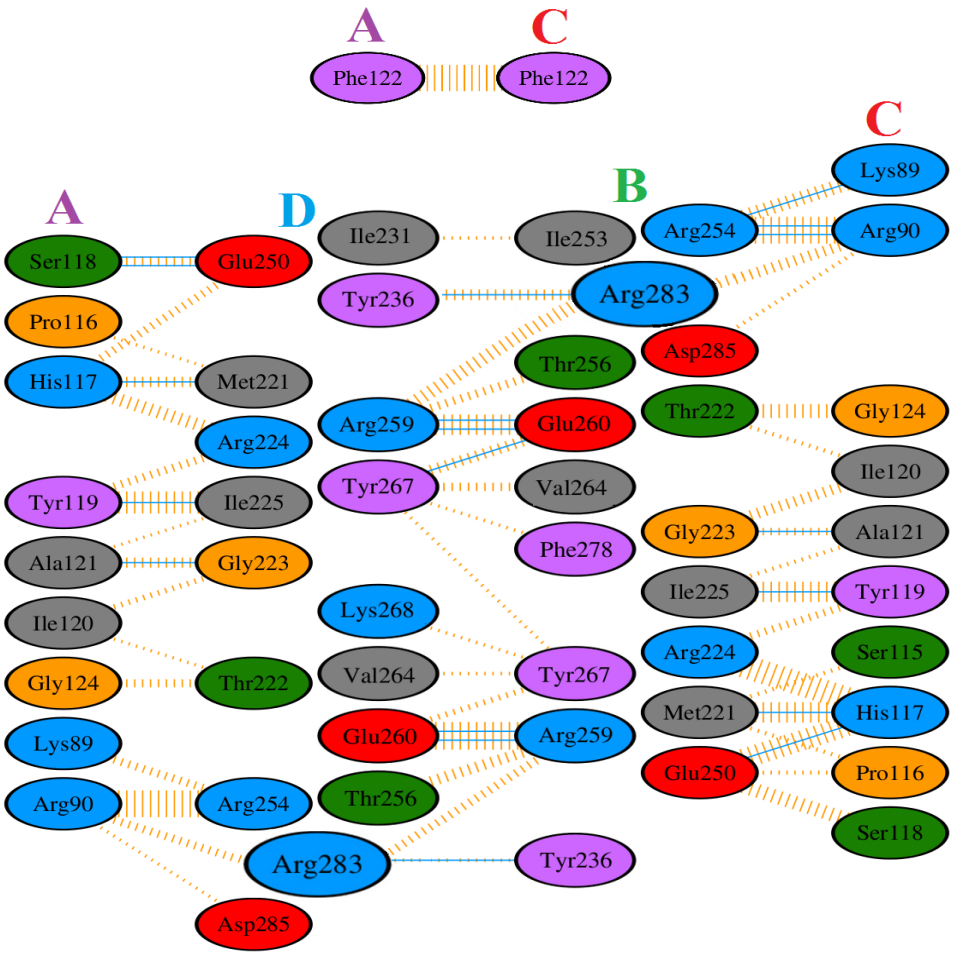
W podjednostce α heterodimeru αβ występuje pięć helis α (αH2, αH3, αH4, αH5, αH7) oraz jedna helisa 3_{10} (αH6), a w podjednostce β są trzy helisy α (βH1, βH2, βH3).

Mniejszy arkusz β złożony jest z czterech antyrównoległych nici β należących wyłącznie do podjednostki β (βS7, βS6, βS5, βS4). Większy arkusz złożony jest z ośmiu głównie antyrównoległych nici β należących do obu podjednostek (βS9, βS8, αS1, βS1, βS2, αS2, αS3, αS4) i uzupełniony jest dziewiątą nicią β należącą do podjednostki β sąsiedniego heterodimeru (zielony element βS3). Większy arkusz β spina zatem cały heterodimer αβ oraz łączy go z sąsiednim heterodimerem tworząc spójną całość.

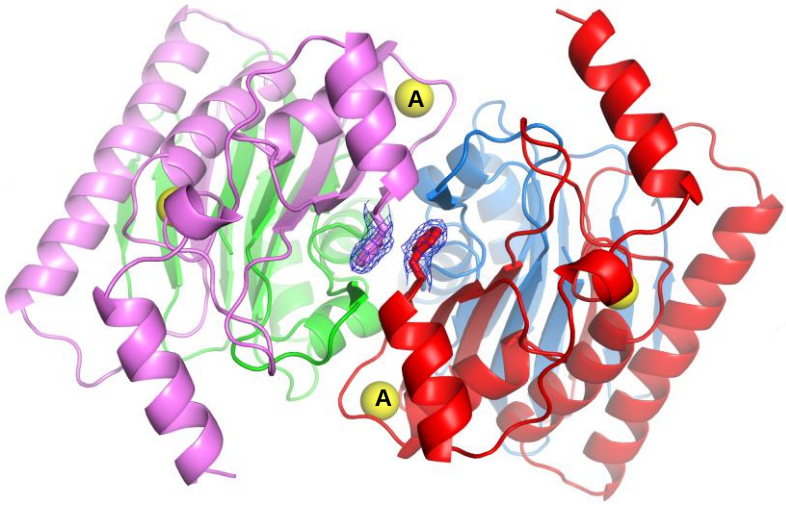
Czwartorzędowa struktura PvAspG1 stabilizowana jest przez wiele różnorakich oddziaływań występujących pomiędzy wszystkimi łańcuchami białkowymi na zasadzie każdy z każdym. Oddziaływania te (poza najliczniejszymi występującymi między podjednostkami α i β tego samego heterodimeru (AB) i (CD)) podsumowałam na Rysunku 7.4. Jedną z ciekawszych interakcji jest oddziaływanie pomiędzy dwoma łańcuchami bocznymi reszt aminokwasowych Phe122 z łańcuchów białkowych A i C (Rysunek 7.5). Reszty te tworzą jedyny kontakt podjednostek α dwóch heterodimerów i mają bardzo nietypową konformację (występują na granicy obszarów dozwolonych wykresu Ramachandrana), przy czym ich doskonała gęstość elektronowa nie pozostawia wątpliwości co do ich konformacji.

W strukturze PvAspG1 istnieją cztery miejsca, w których związane są jony metali (S_A, S_C, A_A, A_C) (Rysunki 7.2, 7.3 i 7.5). Dwa z nich (S_A i A_A) zlokalizowane są w łańcuchu A (w podjednostce α heterodimeru AB), a kolejne dwa (S_C i A_C) znajdują się w analogicznych miejscach tworzonych przez łańcuch C (w podjednostce α heterodimeru CD). Wszystkie miejsca usytuowane są w pętlach na zakończeniach helis. Pętle wiążące jony metali alkaicznych w miejscach S_A i S_C zostały już zidentyfikowane w strukturach asparaginaz potasowo-niezależnych jako pętle wiążące sód, jednak w niniejszej pracy zmieniłam ich nazwę na pętle stabilizujące, gdyż można w nich znaleźć zarówno sód jak i potas (Rozdział 7.2.2.4). Znajdują się one pomiędzy helisą αH3 i nicią β αS2. Pętle wokół miejsc A_A i A_C nie zostały dotąd zidentyfikowane w strukturach asparaginaz potasowo-niezależnych. Nazwałam je pętlami aktywującymi (Rozdział 7.2.2.5). Pętle te znajdują się pomiędzy helisą αH4 i nicią β αS4. Opisane miejsca występowania metali (S_A, S_C, A_A, A_C) koordynują różne jony w różnych modelach asparaginazy. Model PvAspG1-K zawiera we wszystkich czterech pozycjach jony potasu, PvAspG1-Na zawiera cztery jony sodu, a w strukturze PvAspG1-K/Na występują cztery pary jonów K^+/Na^+ o częściowych obsadzeniach.

Rysunek 7.4. Odziaływania pomiędzy resztami aminokwasowymi czterech łańcuchów struktury PvAspG1. Schemat wygenerowany na przykładzie PvAspG1-K za pomocą PDBsum (**Laskowski, 2009**). Pominęłam oczywiste wiązania pomiędzy podjednostkami α i β heterodimerów AB i CB. Niebieskie linie ciągłe oznaczają wiązania wodorowe, pomarańczowe przerywane linie to inne kontakty, a ich szerokość jest proporcjonalna do ilości tych oddziaływań.



Rysunek 7.5. Odziaływanie pomiędzy resztami aminokwasowymi dwóch podjednostek α struktury PvAspG1. Struktura pokazana jest wzdłuż niekrystalograficzej osi dwukrotnej. Oddziałujące reszty Phe122A i Phe122C uwidocznione są z mapą gęstości elektronowej 2*Fo-Fc* (1.5 σ). Łańcuchy polipeptydowe oznaczyłam kolorami: A - fioletowy (podjednostka α), B - zielony (β), C - czerwony (α), D - niebieski (β); żółte kule reprezentują jony metalu w pozycjach A i S.

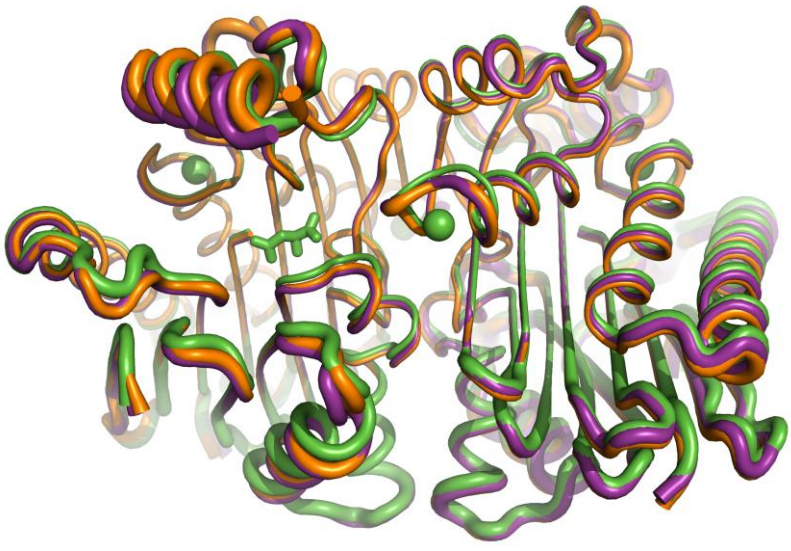


W strukturze asparaginazy potasowo-zależnej z fasoli zwyczajnej, ze względu na jej dimeryczny charakter znajdują się dwa miejsca aktywne, po jednym w każdej podjednostce β hetrodimerów AB i CD. Te dwa równoważne miejsca aktywne PvAspG1 położone są w pozycjach analogicznych do miejsc aktywnych innych asparaginaz typu roślinnego. Lokalizację miejsca aktywnego wyznacza N-końcowa reszta aminokwasowa podjednostki β, tj. katalityczna treonina, będąca elementem większego arkusza β położonego w rdzeniu struktury. Miejsce aktywne danego heterodimeru (np. AB) jest usytuowane pomiędzy pętlą stabilizująca (pętla wokół miejsca S_A) tego samego heterodimeru (AB) i pętlą aktywującą (pętla wokół miejsca A_C) sąsiedniego heterodimeru (CD) (Rysunek 7.2).

7.2.2.2. Sekwencja białka w świetle gęstości elektronowej

Zbadane przeze mnie trzy modele asparaginazy potasowo-zależnej z fasoli zwyczajnej (PvAspG1-K, PvAspG1-Na, PvAspG1-K/Na) są tak samo pofałdowane, mają identyczną topologię heterodimerów αβ i zawierają dwa miejsca aktywne oraz cztery pętle koordynujące jony metali w pozycjach S_A, S_C, A_A, A_C.

Aby określić stopień podobieństwa trzech modeli nałożyłam struktury przy użyciu programu ALIGN (**Satow *et al.*, 1986**). Wartość r.m.s.d. dla pozycji wszystkich atomów C_α poszczególnych par wynosi: 0.44 Å dla PvAspG1-K / PvAspG1-Na, 0.32 Å dla PvAspG1-Na / PvAspG1-K/Na i 0.41 Å dla PvAspG1-K/Na / PvAspG1-K. Wartości te są niskie i świadczą o niemal identycznym zwoju (Rysunek 7.6).



Rysunek. 7.6. Porównanie trzech modeli PvAspG1. Nałożenie atomów C_α modeli PvAspG1-K (zielony), PvAspG1-Na (fioletowy) i PvAspG1-K/Na (pomarańczowy). Zielone kule reprezentujące jony potasu oraz zielona cząsteczka L-Asp reprezentująca ligand związany w miejscu aktywnym pochodzą ze struktury PvAspG1-K. Im wyższa wartość czynnika temperaturowego reszt aminokwasowych tym grubszy łańcuch polipeptydowy na rysunku.

Struktury PvAspG1-K, PvAspG1-Na i PvAspG1-K/Na modelowałam w oparciu o trzy odrębne mapy gęstości elektronowej, dlatego istnieją w nich obszary, w których przebieg łańcuchów głównych lub konformacje łańcuchów bocznych reszt aminokwasowych są nieco zróżnicowane. Rejony te znajdują się najczęściej przy powierzchni, lub nawet wystają poza powierzchnię białka, dlatego są naturalnie bardziej elastyczne, mają wyższe czynniki temperaturowe i są niekiedy zdefiniowane przez fragmentaryczną gęstość elektronową. Jednak znaczna większość niezgodności to minimalne różnice wynikające z prostego skorelowanego przesunięcia względem siebie całych grup atomów tworzących analogiczne elementy strukturalne.

Tabela 7.5. Podsumowanie zawartości trzech modeli PvAspG1.

cecha	łańcuch	PvAspG1-K	PvAspG1-K/Na	PvAspG1-Na
reszty aminokwasowe wymodelowane w mapach gęstości elektronowej	A	2 - 158	2 - 156	2 - 158
	B	196 - 326	196 - 326	196 - 326
	C	2 - 158	2 - 159	2 - 156
	D	196 - 326	196 - 326	196 - 326
jon koordynowany w pętli stabilizującej - $S_{A/C}$ (obsadzenie)	A	K^+ (1.0)	K^+/Na^+ (0.3/0.7)	Na^+ (1.0)
	C	K^+ (1.0)	K^+/Na^+ (0.3/0.7)	Na^+ (1.0)
jon koordynowany w pętli aktywującej - $A_{A/C}$ (obsadzenie)	A	K^+ (1.0)	K^+/Na^+ (0.3/0.7)	Na^+ (1.0)
	C	K^+ (1.0)	K^+/Na^+ (0.3/0.7)	Na^+ (1.0)
cząsteczki obecne w miejscu aktywnym (obsadzenie)	B	L-Asp (0.7)	kilka H_2O	kilka H_2O
	D	kilka H_2O	kilka H_2O	kilka H_2O
liczba cząsteczek wody	ABCD	114	421	286
inne składniki	ABCD	-	21 NO_3^-	-

Z powodu braku fragmentów mapy gęstości elektronowej wymodelowanie niektórych reszt aminokwasowych okazało się niemożliwe (Tabela 7.5). Biorąc pod uwagę sekwencję asparaginazy z fasoli zwyczajnej, podjednostka α powinna zawierać reszty o numerach od 1 do 195, a podjednostka β od 196 do 326. Natomiast uwzględniając konstrukcje wektora ekspresyjnego, na N-końcu podjednostki α należało spodziewać się dwóch dodatkowych reszt aminokwasowych (o numerach -1 i 0), stanowiących pozostałość po procesie produkcji i oczyszczania białka (**Bejger *et al.*, 2014**). We wszystkich trzech modelach PvAspG1 jednoznaczne mapy gęstości elektronowej definiują położenia wszystkich spodziewanych reszt w podjednostkach β. Natomiast w podjednostkach α brakuje fragmentów mapy gęstości elektronowej zarówno dla N- jak i C-końca łańcucha. Wszystkie trzy modele nie zawierają reszt o

numerach -1, 0, 1 z końca aminowego, oraz od trzydziestu sześciu do trzydziestu dziewięciu reszt z końca karboksylowego. Brak gęstości elektronowej dla tych reszt jest prawdopodobnie skutkiem ich nieuporządkowania, które wynika z ich labilności. Jak wykazały kontrole na żelach SDS-PAGE, wszystkie trzy preparaty białka przeznaczone do krystalizacji, a właściwie ich podjednostki α i β posiadały masy odpowiadające teoretycznym masom podjednostek pełnej długości.

7.2.2.3. Architektura miejsca aktywnego

Położenie miejsca aktywnego PvAspG1 wyznacza N-końcowa reszta aminokwasowa podjednostki β - treonina 196, która pełni role katalitycznego nukleofila w reakcji hydrolizy L-asparaginy i jednocześnie jest resztą biorącą udział w autokatalitycznym cięciu łańcuchów polipeptydowych niedojrzałego białka (**Michalska *et al.*, 2008-B**). Miejsce aktywne jest zlokalizowane pomiędzy pętlą stabilizującą tego samego heterodimeru a pętlą aktywującą sąsiedniego heterodimeru (Rysunek 7.2). Znajduje się ono w wydłużonej szczelinie pomiędzy dwoma arkuszami β, które tworzą rdzeń struktury. Krawędzie szczeliny uformowane są głównie przez kilka pętli z obu podjednostek heterodimeru αβ łączących elementy helikalne z łańcuchami β.

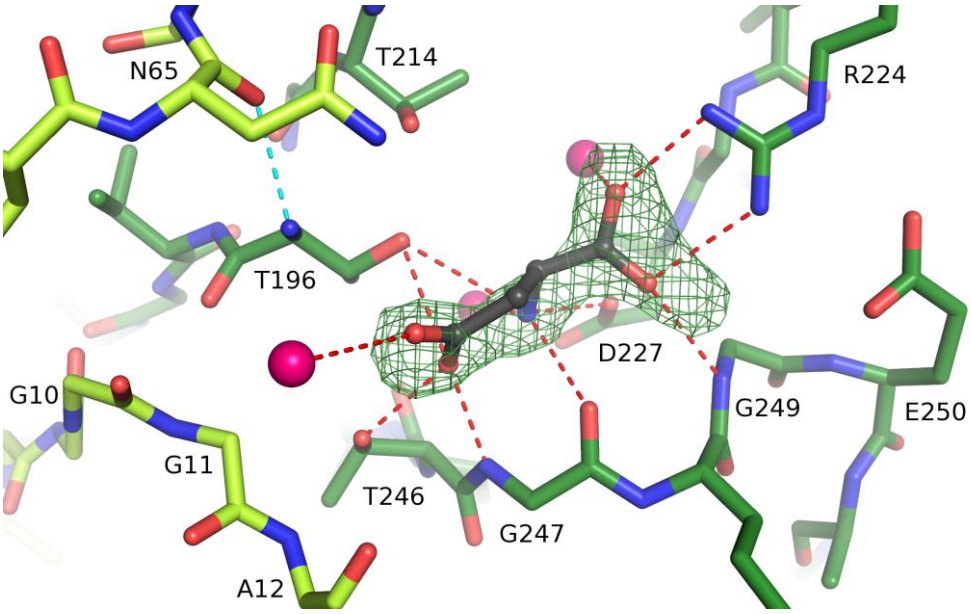
Spośród trzech modeli asparaginazy potasowo-zależnej tylko PvAspG1-K ma jedno z miejsc aktywnych obsadzone L-asparaginianem - produktem reakcji katalizowanej przez ten enzym. Jest to miejsce w heterodimerze AB, a ligand wymodelowałam w niepełnym obsadzeniu (70%). Drugie miejsce aktywne PvAspG1-K (w heterodimerze CD) jest wypełnione kilkoma cząsteczkami wody. Częściowe obsadzenie cząsteczką L-asparaginianu jest prawdopodobnie spowodowane zbyt małym jego nadmiarem dodanym do krystalizacji (nadmiar molowy 1:3). Jednak różnica ta umożliwiła mi porównanie w ramach jednej struktury, a więc w identycznych warunkach, miejsca obsadzonego produktem - L-Asp i pustego. W strukturach PvAspG1-Na i -K/Na miejsca aktywne nie wiążą ligandów, mimo zwiększenia ilości L-asparaginianu dodawanego w trakcie krystalizacji (nadmiar molowy 1:25). Jony sodu obecne w tych kryształach (lub też brak jonów potasu) ewidentnie wpływają na możliwość związania liganda w tych strukturach.

Cząsteczka L-asparaginianu obecna w modelu PvAspG1-K prawie całkowicie wypełnia wnękę katalityczną. Pozostaje jednak sporo wolnej przestrzeni od strony grupy β-karboksylowej liganda, na przedłużeniu jego łańcucha bocznego. Cała

przestrzeń wnęki z pewnością pomieściłaby oba rodzaje substratów asparaginaz, L-asparaginę, jak również peptyd izoaspartylowy. Odcinana w trakcie reakcji reszta aminokwasowa β-peptydu mogłaby oddziaływać ze znanym już z asparaginaz potasowo-niezależnych (**Michalska i Jaskólski, 2006**) motywem GAG (Gly11-Ala12-Gly13), znajdującym się w obszarze bogatym w Gly tworzonym przez fragment łańcucha podjednostki α heterodimeru AB.

L-Asparaginian jest stabilizowany we wnęce katalitycznej za pomocą wiązań wodorowych utworzonych z pięcioma resztami aminokwasowymi i trzema cząsteczkami wody (Rysunek 7.7, Tabela 7.6). Wspomniane reszty pochodzą tylko z podjednostki β heterodimeru AB. Grupa α-karboksylowa liganda tworzy wiązanie z grupą amidową Gly249 oraz dwa wiązania w mostku solnym z grupą guanidynową łańcucha bocznego Arg224. Grupa α-aminowa oddziałuje z karbonylem Gly247 oraz łańcuchem bocznym Asp227. Z kolei grupa β-karboksylowa oddziałuje z atomem azotu łańcucha głównego Gly247 oraz grupą hydroksylową łańcucha bocznego Thr246. Te dwie ostatnie reszty są prawdopodobnie odpowiedzialne za wytworzenie tzw. wnęki oksyanionowej, która stabilizuje ujemny ładunek na atomie tlenu substratu podczas reakcji hydrolizy. Wszystkie wymienione wiązania między L-Asp i białkiem występują także w asparaginazach K^+-niezależnych (**Michalska *et al.*, 2005; Nomme *et al.*, 2012**).

Rysunek 7.7. Miejsce aktywne PvAspG1-K (hetreodimer AB) z wymodelowanym ligandem L-Asp. Łańcuch białkowy A - jasnozielony, B – ciemnozielony, ligand - czarny, atomy O - czerwone, N - niebieskie, H_2O - różowe kule. Ligand pokazałam w dodatniej gęstości różnicowej F_o-F_c (2.5 σ) ze wszystkimi oddziaływaniami, jakie tworzy z białkiem (czerwone przerywane linie). Na rysunku zaznaczyłam także stabilizujące wiązanie wodorowe pomiędzy katalityczną resztą Thr196 i resztą Asn65 (błękitne przerywane linie).



Interesujące wyniki uzyskałam porównując wszystkie sześć miejsc aktywnych z trzech struktur asparaginazy z fasoli zwyczajnej. Okazuje się, że reszty aminokwasowe tworzące ekwiwalentne miejsca aktywne struktury PvAspG1-K (z heterodimeru AB - obsadzone L-Asp, i heterodimeru CD - puste) mają identyczne konformacje. Zatem można stwierdzić, że obecność/nieobecność liganda nie ma wpływu na budowę miejsca aktywnego. Co ciekawe, nie zawierające liganda cztery miejsca aktywne modeli PvAspG1-Na i PvAspG1-K/Na różnią się od miejsc aktywnych modelu PvAspG1-K. Różnica polega przede wszystkim na odmiennej konformacji łańcucha bocznego Arg224 (Rozdział 7.2.2.6).

Tabela 7.6. Wiązania wodorowe występujące w miejscu aktywnym PvAspG1-K (heterodimer AB).

	atomy oddziałujące		odległość (Å)
ligand L-asparaginian	$N\cdots O_{\delta1}$	Asp227	2.8
	$N\cdots O$	H_2O	2.9
	$N\cdots O$	Gly247	3.2
	$N\cdots O_{\gamma1}$	Thr196	3.3
	$O\cdots N_{\eta1}$	Arg224	2.8
	$O\cdots O$	H_2O	3.1
	$OXT\cdots N$	Gly249	2.8
	$OXT\cdots N_{\eta2}$	Arg224	3.0
	$O_{\delta1}\cdots O$	H_2O	2.9
	$O_{\delta2}\cdots N$	Gly247	2.5
	$O_{\delta2}\cdots O_{\gamma1}$	Thr246	2.6
	$O_{\delta2}\cdots O_{\gamma1}$	Thr196	3.0
Thr196	$O_{\gamma1}\cdots N$	Thr196	2.9
	$O_{\gamma1}\cdots O_{\gamma1}$	Thr214	2.7
	$N\cdots O$	Asn65	2.8
	$N\cdots O$	H_2O	3.0
	$O\cdots O$	H_2O	3.5
	$O\cdots N$	Gly11	2.9
	$O\cdots O$	Gly10	3.3
Arg224	$N_{\eta1}\cdots O$	H_2O	3.0
	$N_{\eta2}\cdots O_{\varepsilon2}$	Glu250	3.3
	$N_{\varepsilon}\cdots O_{\varepsilon1}$	Glu250	3.1
	$N_{\varepsilon}\cdots O$	H_2O	3.4

7.2.2.4. Pętla stabilizująca

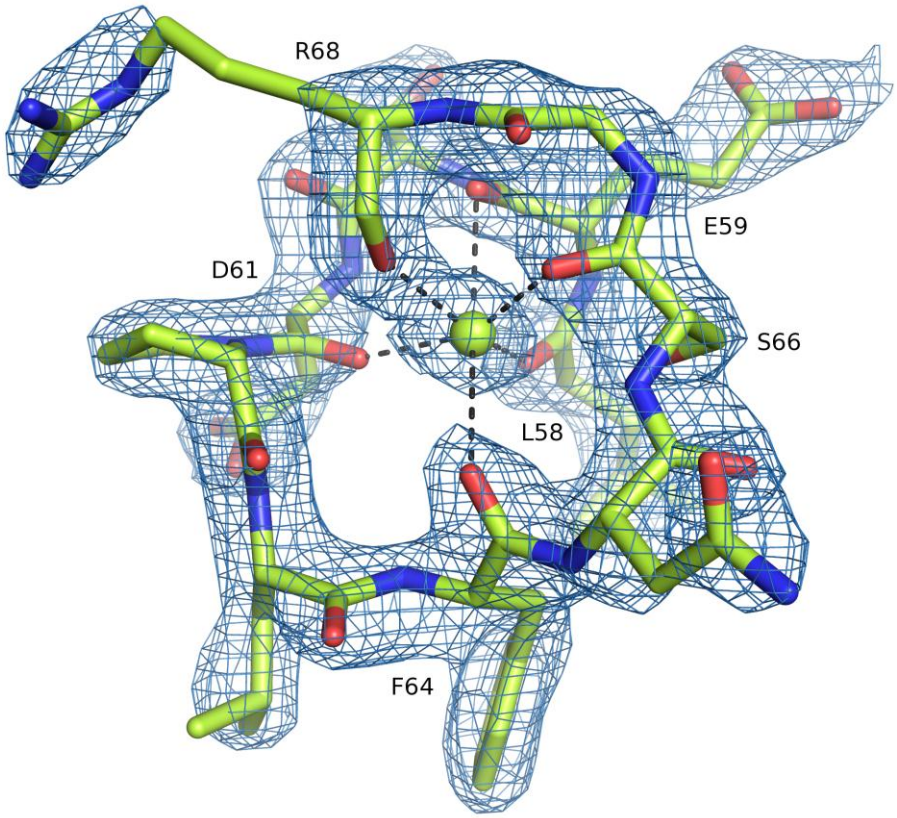
Dwie analogiczne pętle stabilizujące, zawierające jony położone w miejscach S_A i S_C, utworzone są odpowiednio przez łańcuchy A i C struktury PvAspG1 (tj. przez

podjednostki α heterodimerów). Pętla stabilizująca znajduje się na zakończeniu helisy αH3 i jest tworzona przez 11 kolejnych reszt aminokwasowych (Leu58 - Arg68) okrążających półtora raza centralnie koordynowany jon metalu. Sfera koordynacyjna jonu metalu ma geometrię oktaedryczną, a wszystkie wiązania (jon metalu - tlen) utworzone są wyłącznie przez grupy karbonylowe łańcucha głównego reszt tworzących pętlę, podczas gdy łańcuchy boczne tych reszt są skierowane na zewnątrz pętli (Rysunek 7.8).

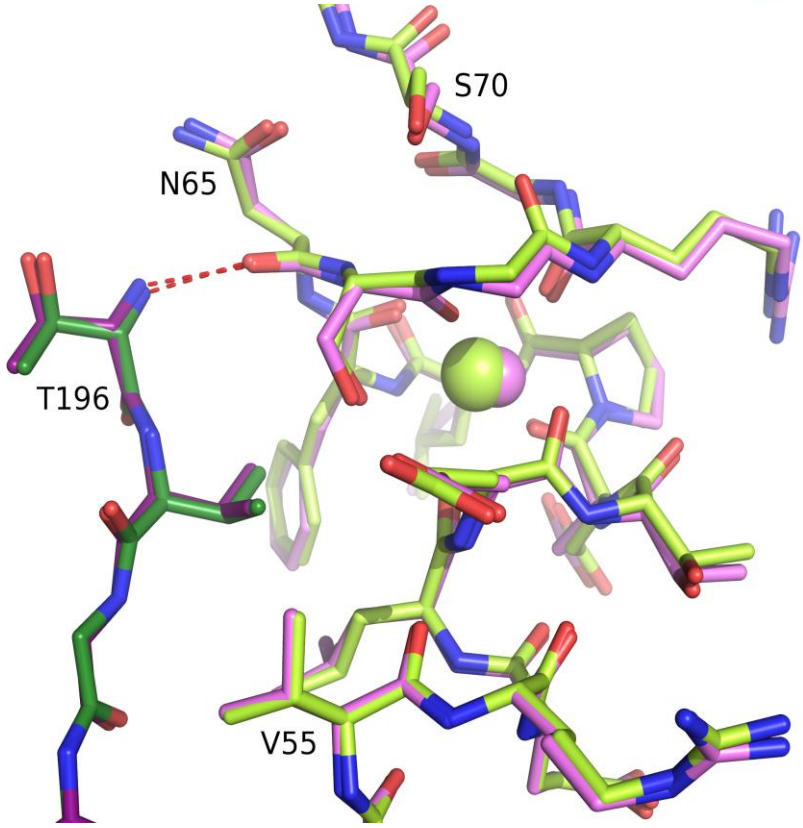
Ten rodzaj pętli nazwałam stabilizująca ze względu na funkcję jaką pełni. Stabilizuje ona miejsce aktywne za pomocą wiązania wodorowego pomiędzy wolną grupą α-aminową katalitycznej treoniny 196, a karbonylem łańcucha głównego jednej z reszt tworzących pętle - Asn65 (Rysunki 7.7 i 7.9). Oddziaływanie to prawdopodobnie wpływa na właściwą orientacje Thr196, niezbędną dla ataku nukleofilowego na grupę β-amidową substratu, choć niewykluczone, że pełni jeszcze inną ważną role (Rozdział 7.3.3).

W strukturze PvAspG1-K obie pętle stabilizujące są obsadzone jonami potasu. Oba jony K^+ posiadają sześć wiązań koordynacyjnych o geometrii oktaedrycznej (Tabela 7.7). W PvAspG1-Na te same pętle zawierają jony sodu, które również posiadają sześć wiązań o geometrii oktaedrycznej (Tabela 7.7). W PvAspG1-K/Na obsadzenie obu pętli stabilizujących jest mieszane - zawierają one w 30% potas i 70% sód. Z tego względu geometria sfery koordynacyjnej jak i liczba tworzonych wiązań jest bardziej nieregularna (Tabela 7.8).

Pętle stabilizujące we wszystkich trzech modelach PvAspG1 są bardzo podobne, zarówno co do przebiegu łańcucha głównego jak i konformacji łańcuchów bocznych (Rysunek 7.9). Fakt ten potwierdzają niskie wartości r.m.s.d. wyznaczone dla wszystkich atomów z obszaru pętli pomiędzy resztami Val55 i Ala71 (Tabela 7.9). Jedynie pozycje jonów z różnych struktur są względem siebie nieco przesunięte. Jony potasu z PvAspG1-K są oddalone od jonów sodu z PvAspG1-Na o ok. 0.5 Å, natomiast odległość pomiędzy jonami K^+ i Na^+ w strukturze PvAspG1-K/Na osiąga nawet ok. 1.9 Å. Mogę zatem wnioskować, że architektura pętli stabilizujących pozwala im skoordynować dostępne w otoczeniu jony metali, przy jednoczesnym zachowaniu funkcji stabilizującej.



Rysunek 7.8. Pętla stabilizująca struktury PvAspG1-K (heterodimer AB). Poszczególne atomy łańcucha polipeptydowego rozróżniają kolory: C - jasnozielony, O - czerwony, N – niebieski. Pętle pokazałam w gęstości elektronowej $2F_o$-F_c (2.0 σ). Jasno zielona kula reprezentuje jon potasu, czarne linie przerywane to wiązania koordynacyjne.



Rysunek 7.9. Nałożenie atomów C_α obszaru pętli stabilizującej dwóch modeli PvAspG1 (heterodimery AB). Model PvAspG1-K reprezentują jasnozielone reszty aminokwasowe (łańcuch A), ciemnozielone reszty (łańcuch B) oraz jasnozielona kula (jon potasu). Model PvAspG1-Na reprezentują jasnofioletowe reszty aminokwasowe (łańcuch A), ciemnofioletowe reszty (łańcuch B) oraz jasnofioletowa kula (jon sodu). Dodatkowo zaznaczyłam atomy O na czerwono, N na niebiesko oraz wiązanie wodorowe stabilizujące miejsce aktywne (katalityczną Thr196) jako czerwone linie przerywane.

Tabela 7.7. Koordynacja jonu metalu w pętlach stabilizujących struktur PvAspG1-K i PvAspG1-Na. Wiązania spełniają warunki stereochemiczne: mają oktaedryczną geometrię i są najkrótszymi z oddziaływań między jonem metalu i białkiem. W nawiasach podaję czynniki temperaturowe ($Å^2$) dla danego atomu/jonu metalu, odległości podaję w Å.

	PvAspG1-K				**PvAspG1-Na**			
	S_A : K^+ (41.7)		S_C : K^+ (57.8)		S_A : Na^+ (40.2)		S_C : Na^+ (40.8)	
reszta amino-kwasowa	atom	odległość	atom	odległość	atom	odległość	atom	odległość
Leu58	O (37.1)	2.93	O (47.7)	2.85	O (39.2)	2.96	O (36.7)	2.84
Glu59	O (29.4)	2.95	O (47.8)	2.86	O (38.4)	2.60	O (40.5)	2.55
Asp61	O (42.0)	2.52	O (50.6)	2.58	O (31.1)	2.27	O (33.4)	2.27
Phe64	O (29.5)	3.10	O (49.9)	3.19	O (39.6)	3.32	O (33.6)	3.20
Ser66	O (25.4)	2.64	O (43.4)	2.62	O (31.5)	2.35	O (33.3)	2.23
Arg68	O (31.1)	2.56	O (36.1)	2.66	O (30.1)	2.23	O (32.1)	2.33

Tabela 7.8. Koordynacja jonu metalu w pętlach stabilizujących struktury PvAspG1-K/Na. Wymienione zostały wszystkie oddziaływania do 3.6 Å, choć nie wszystkie można uznać za wiązania. W nawiasach podaję czynniki temperaturowe ($Å^2$) dla danego atomu/jonu metalu, odległości podaję w Å. Wartości procentowe dotyczą współczynników obsadzenia.

PvAspG1-K/Na	S_A :	K^+ 30% (29.7)	Na^+ 70% (18.7)	S_C :	K^+ 30% (26.8)	Na^+ 70% (13.2)
reszta aminokwasowa	atom	odległość	odległość	atom	odległość	odległość
Leu58	O (23.4)	2.69	3.22	O (20.5)	2.53	3.34
Glu59	O (23.7)	-	2.39	O (21.3)	-	2.33
Asp61	O (21.1)	2.86	2.27	O (17.5)	2.93	2.33
Phe64	O (23.1)	2.68	-	O (21.6)	2.83	-
Ser66	O (21.7)	2.71	2.29	O (20.5)	2.76	2.24
	Oγ (29.7)	2.97	-	Oγ (26.1)	3.10	-
Arg68	O (24.6)	-	2.20	O (19.8)	-	2.16
Woda	O (28.1)	2.64	-	O (24.5)	2.59	-

Tabela 7.9. Porównanie pętli stabilizujących modeli PvAspG1-K, PvAspG1-Na i PvAspG1-K/Na. Nałożyłam wszystkie atomy reszt 55-71 w programie ALIGN (**Satow *et al.*, 1986**). Parametry r.m.s.d. podaję w Å.

		PvAspG1-K		PvAspG1-Na		PvAspG1-K/Na	
		S_A	S_C	S_A	S_C	S_A	S_C
PvAspG1-K	S_A	0	0.16	0.22	0.18	0.18	0.18
	S_C		0	0.23	0.20	0.18	0.19
PvAspG1-Na	S_A			0	0.11	0.16	0.14
	S_C				0	0.13	0.14
PvAspG1-K/Na	S_A					0	0.09
	S_C						0

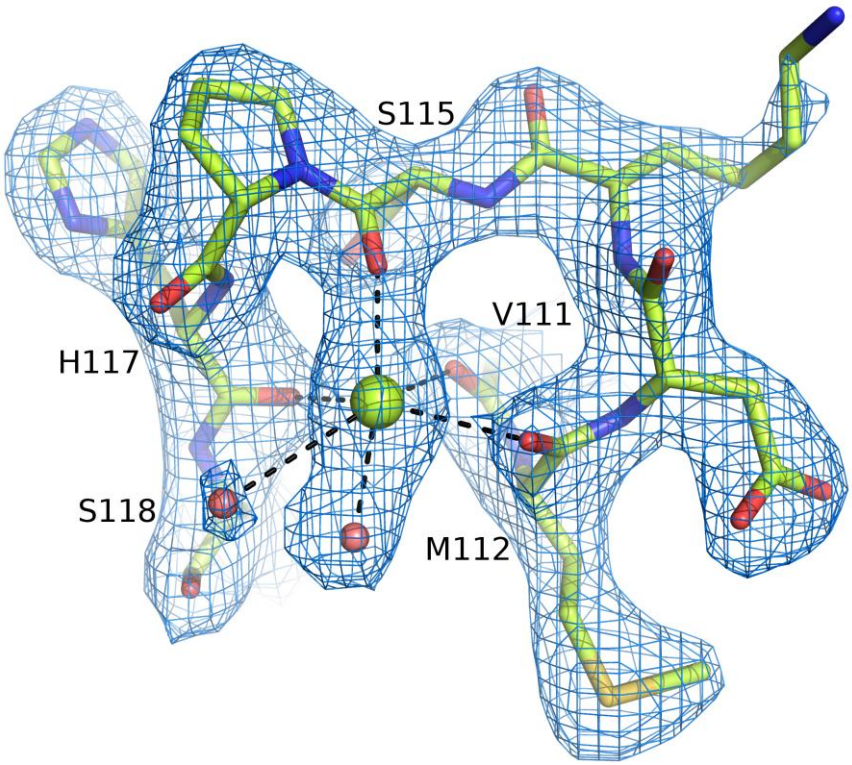
7.2.2.5. Pętla aktywująca

Pętla aktywująca jest charakterystyczną cechą asparaginaz potasowo-zależnych, ponieważ nie występuje w takiej roli w żadnej ze znanych struktur enzymów potasowo-niezależnych (LlA, EcAIII, hASNase3, gpASNase3). W strukturze PvAspG1 istnieją dwie analogiczne pętle aktywujące tworzone przez łańcuchy białkowe A (pętla wiążąca jon metalu w pozycji A_A) i C (pętla wiążąca jon metalu w pozycji A_C). Pętle te znajdują się na zakończeniach helisy αH1 (podjednostek α heterodimerów AB i CD). Pętla aktywująca zbudowana jest z łańcucha ośmiu kolejnych reszt aminokwasowych (Val111 - Ser118) wykonującego jeden pełen obrót wokół centralnie położonego jonu metalu. Sfera koordynacyjna jonu metalu ma geometrię oktaedryczną i tworzona jest przez cztery grupy karbonylowe łańcucha głównego i dwie cząsteczki wody, które stabilizowane są przez łańcuch boczny Ser118, (ostatniej reszty tworzącej pętle) (Rysunek 7.10).

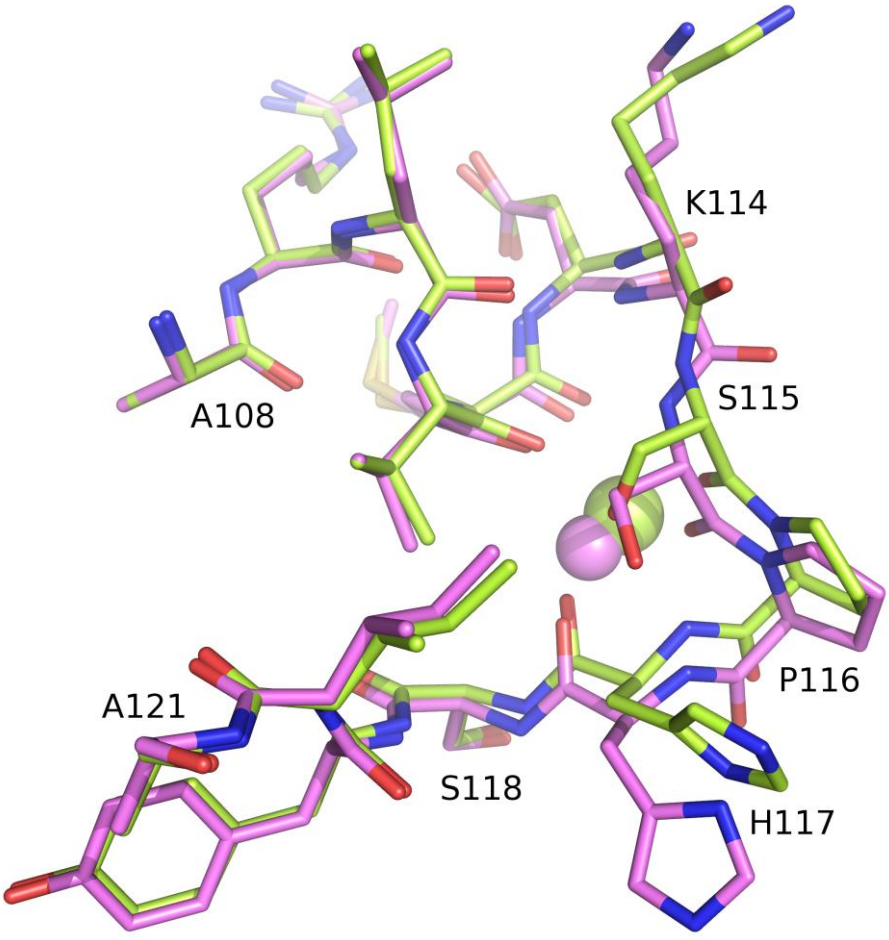
Należy zauważyć, że taka geometria sfery koordynacyjnej jest zachowana w przypadku modeli PvAspG1-K i PvAspG1-Na, gdzie pętle aktywujące obsadzone są odpowiednio przez jony potasu lub sodu (Tabela 7.10), jednak w przypadku PvAspG1-K/Na układ tworzonych wiązań jest bardziej skomplikowany ze względu na podwójne obsadzenie każdej z dwóch pętli (Tabela 7.11).

Aby porównać ukształtowanie pętli aktywujących z trzech modeli asparaginazy z fasoli zwyczajnej nałożyłam w programie ALIGN (**Satow *et al.*, 1986**) wszystkie atomy obszaru pętli obejmującego reszty Ala108 - Ala121 (Tabela 7.12). Wyniki wskazują na znaczne podobieństwo rozmieszczenia przestrzennego atomów w pętlach PvAspG1-Na i PvAspG1-K/Na oraz wyraźną różnicę pomiędzy PvAspG1-K i PvAspG1-Na.

Aby dokładniej scharakteryzować różnice w konformacji pętli aktywującej PvAspG1-K i PvAspG1-Na, nałożyłam wszystkie atomy C_α tych struktur wykorzystując narzędzie SSM (**Krissinel i Henrick, 2004**) z programu COOT (**Emsley i Cowtan, 2004**). Obie pętle aktywujące PvAspG1-K są obsadzone jonami potasu, a obie pętle PvAspG1-Na jonami sodu, przy czym względne pozycje jonów K^+ i Na^+ są oddalone od siebie o ok. 1.0 Å. Wyraźne różnice widać także w rozmieszczeniu reszt aminokwasowych tworzących krzywiznę pętli, zmieniony jest zarówno przebieg łańcucha głównego (między resztami Asp113 - Ser118) jak i konformacje grup bocznych czterech reszt (Lys114, Ser115, Pro116, His117) (Rysunek 7.11). Co ciekawe,



Rysunek 7.10. Pętla aktywująca struktury PvAspG1-K (heterodimer CD). Poszczególne atomy łańcucha polipeptydowego rozróżniają kolory: C - jasnozielony, O - czerwony, N – niebieski, S - żółty. Pętlę pokazałam w gęstości elektronowej $2F_o$-F_c (2.0 σ). Jasno zielona kula reprezentuje jon potasu, czarne linie przerywane to wiązania koordynacyjne.



Rysunek 7.11. Nałożenie atomów C_α obszaru pętli aktywującej dwóch modeli PvAspG1 (heterodimery CD). Model PvAspG1-K reprezentują jasnozielone reszty aminokwasowe (łańcuch C) oraz jasnozielona kula (jon potasu). Model PvAspG1-Na reprezentują jasnofioletowe reszty aminokwasowe (łańcuch C) oraz jasnofioletowa kula (jon sodu). Dodatkowo zaznaczyłam atomy O na czerwono, N na niebiesko, S na żółto.

Tabela 7.10. Koordynacja jonu metalu w pętlach aktywujących struktur PvAspG1-K i PvAspG1-Na. W nawiasach podaję czynniki temperaturowe (Å^2) dla danego atomu/jonu metalu, odległości podaję w Å.

reszta amino-kwasowa	**PvAspG1-K**				**PvAspG1-Na**			
	A_A : K^+ (44.9)		A_C : K^+ (47.8)		A_A : Na^+ (44.4)		A_C : Na^+ (43.9)	
	atom	odległość	atom	odległość	atom	odległość	atom	odległość
Val111	O (37.2)	2.97	O (36.6)	2.87	O (27.9)	2.69	O (32.3)	2.70
Met112	O (41.9)	2.89	O (37.3)	2.80	O (36.7)	3.38	O (42.8)	3.10
Ser115	O (39.7)	2.65	O (34.2)	2.45	O (41.7)	2.44	O (42.9)	2.40
His117	O (42.3)	2.70	O (36.2)	2.67	O (36.2)	2.34	O (37.4)	2.61
Woda	O (33.6)	2.55	O (28.1)	2.46	O (36.7)	2.23	O (35.1)	2.28
Woda	O (55.6)	2.99	O (41.9)	3.42	O (34.5)	2.34	O (49.4)	3.21

Tabela 7.11. Koordynacja jonu metalu w pętlach aktywujących struktury PvAspG1-K/Na. W nawiasach podaję czynniki temperaturowe (Å^2) dla danego atomu/jonu metalu, odległości podaję w Å. Wartości procentowe dotyczą współczynników obsadzenia.

PvAspG1-K/Na	A_A :	K^+ 30% (34.5)	Na^+ 70% (28.1)	A_C :	K^+ 30% (47.8)	Na^+ 70% (25.0)
reszta aminokwasowa	atom	odległość	odległość	atom	odległość	odległość
Val111	O (19.2)	2.86	2.31	O (18.2)	-	2.43
Met112	O (20.9)	3.41	2.78	O (26.0)	-	3.01
Ser115	O (35.7)	3.23	2.23	O (40.1)	2.73	2.35
His117	O (25.8)	2.36	2.81	O (25.3)	-	2.50
Woda	O (35.4)	2.77	2.39	O (26.1)	3.04	2.63
Woda	O (40.7)	2.79	-	O (51.5)	-	2.55
Woda	O (18.7)	-	2.92	O (45.4)	3.22	-
Woda	-	-	-	O (44.5)	2.68	-

Tabela 7.12. Porównanie obszarów pętli aktywujących modeli PvAspG1-K, PvAspG1-Na i PvAspG1-K/Na. Nałożyłam wszystkie atomy reszt 108-121 w programie ALIGN (**Satow *et al.*, 1986**). Parametry r.m.s.d. podane są w Å.

		PvAspG1-K		PvAspG1-Na		PvAspG1-K/Na	
		A_A	A_C	A_A	A_C	A_A	A_C
PvAspG1-K	A_A	0	0.18	0.40	0.48	0.27	0.26
	A_C		0	0.36	0.38	0.26	0.27
PvAspG1-Na	A_A			0	0.16	0.17	0.15
	A_C				0	0.16	0.17
PvAspG1-K/Na	A_A					0	0.10
	A_C						0

poza samą krzywizną pętli, reszty od strony N-końca (od Asp113) i C-końca (od Ser118) w obu strukturach PvAspG1-K i PvAspG1-Na pozostają dokładnie w tych samych pozycjach i mają identyczne konformacje. Porównanie to wskazuje, że rodzaj związanego przez pętlę aktywującą jonu metalu wpływa na jej konformację, a konkretnie na ułożenie reszt aminokwasowych pomiędzy Asp113 - Ser118.

Konformacja pętli aktywujących PvAspG1-K/Na jest zasadniczo podobna do tych w PvAspG1-Na. Zarówno przebieg łańcucha głównego jak i konformacje większości reszt aminokwasowych są takie same. Wyjątkiem jest reszta His117, której łańcuch boczny przyjmuje dwie różne konformacje. Jedna jest identyczna z konformacją występującą w PvAspG1-K (jej obsadzenie w PvAspG1-K/Na wynosi 30%), druga - taka sama jak w PvAspG1-Na (o obsadzeniu 70%), co w pełni koreluje z podwójnym obsadzeniem tych pętli przez jony K^+ (30%) i Na^+ (70%). Odległość między jonami w pętlach PvAspG1-K/Na wynosi ok. 1.5 Å.

Te nowo odkryte, charakterystyczne pętle nazwałam aktywującymi, ponieważ aktywują one mechanizm przełącznika katalitycznego asparaginazy potasowo-zależnej, opisany w następnym rozdziale.

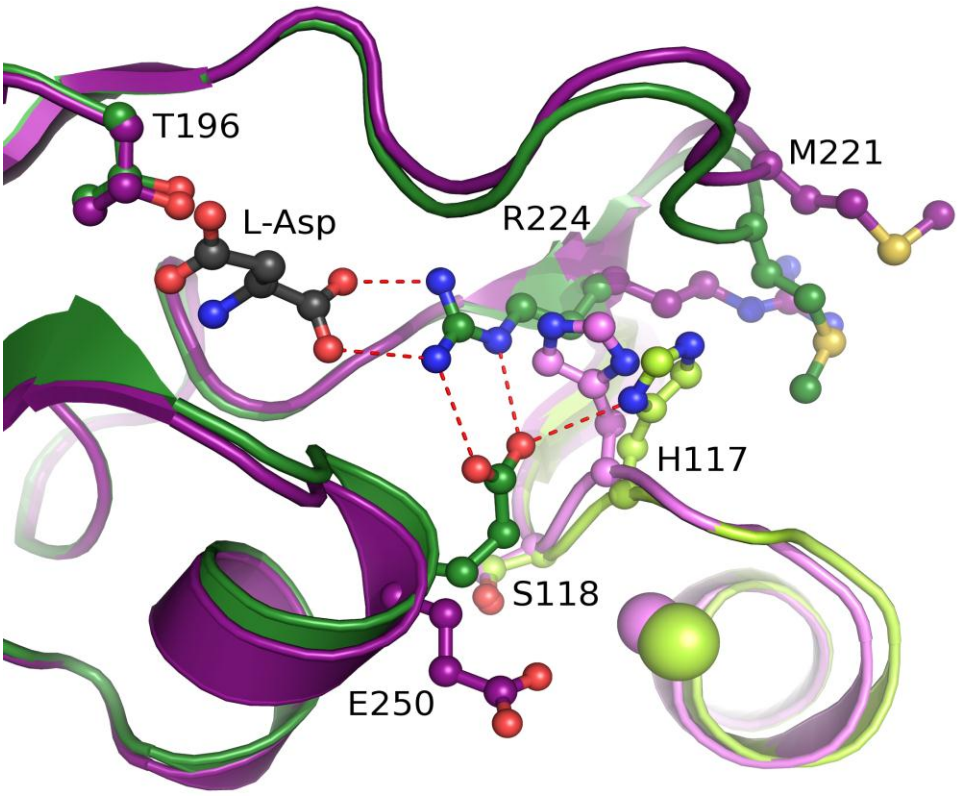
7.2.2.6. Przełącznik katalityczny

Miejsce aktywne PvAspG1 znajduje się pomiędzy pętlą stabilizującą i aktywującą, jednak skoordynowane w tych pętlach jony metali są zbyt oddalone od katalitycznej reszty Thr196, aby wpływać na miejsce aktywne w sposób bezpośredni. Odległości: jon metalu w pętli stabilizującej - grupa OH Thr196 i grupa OH Thr196 - jon metalu z pętli aktywującej wynoszą odpowiednio ~10 Å i ~18 Å. Jednak tak jak jon obecny w pętli stabilizującej usztywnia ją, co pozwala poprzez Asn65 ustabilizować resztę nukleofilową Thr196, tak jon kontrolujący ukształtowanie pętli aktywującej kontaktuje się z miejscem aktywnym poprzez system oddziaływań o charakterze przełącznika katalitycznego.

W strukturze potasowo-zależnej asparaginazy z fasoli zwyczajnej istnieją dwa miejsca aktywne, zatem i dwa przełączniki katalityczne. Przełącznik AD jest pośrednikiem oddziaływania pętli aktywującej zbudowanej z łańcucha aminokwasowego A na miejsce aktywne osadzone w łańcuchu D. Przełącznik CB jest pośrednikiem oddziaływania pętli aktywującej zbudowanej z łańcucha aminokwasowego C na miejsce aktywne osadzone w łańcuchu B. Przełączniki te składają się z trzech reszt

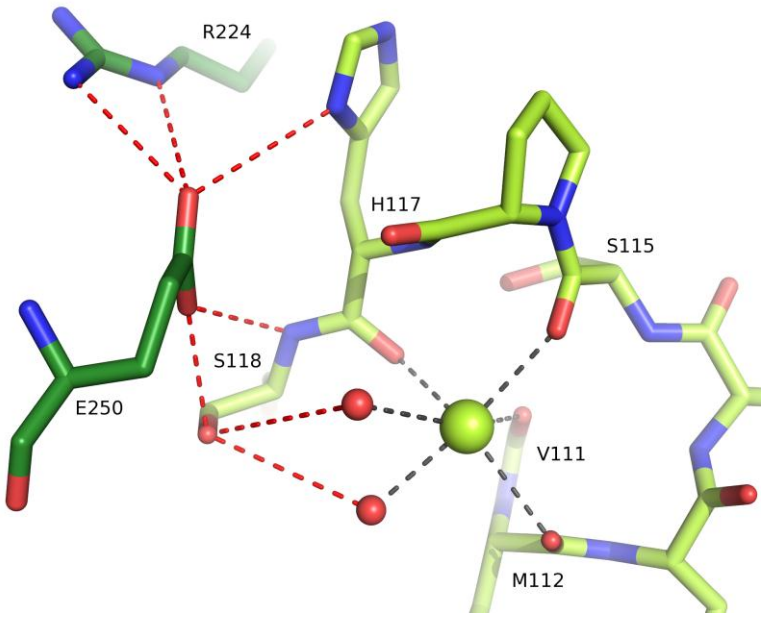
aminokwasowych: His117, Arg224 i Glu250, które przyjmując różne konformacje włączają i wyłączają (maksymalną) aktywność asparaginazową enzymu.

Rysunek 7.12. Zbliżenie obszaru przełącznika katalitycznego obejmujące miejsce aktywne i pętle aktywującą w stanie ON (zielony) i OFF (fioletowy). Nałożenie przełącznika CB struktury PvAspG1-K (łańcuchy polipeptydowe: C - jasnozielony, B - ciemnozielony) z jonem potasu (jasnozielona kula) i ligandem (czarny) na przełącznik CB struktury PvAspG1-Na (C - jasnofioletowy, B - ciemnofioletowy) z jonem sodu (jasnofioletowa kula). Łańcuchy boczne istotnych reszt aminokwasowych zostały pokazane i opisane. Wyróżniono atomy: N - niebieskie, O - czerwone, S - żółte. Włączony przełącznik tworzy wiązania wodorowe zaznaczone jako przerywane czerwone linie.



Stan „włączony” (ON) asparaginazy jest przyjmowany gdy pętla aktywująca koordynuje jon potasu. Stan ten zobrazuję na przykładzie modelu PvAspG1-K, a konkretnie przełącznika CB oddziałującego z miejscem aktywnym (w łańcuchu B), w którym związany jest L-Asp (Rysunek 7.12). Jon potasu związany w pętli aktywującej (łańcuch C) narzuca jej ukształtowanie, a więc i konformację His117 - jednej z reszt pętli, która jest jednocześnie centralnym elementem przełącznika. Pierścień imidazolowy His117 (łańcuch C) jest skierowany w stronę powierzchni białka, co umożliwia znajdującej się głębiej reszcie Arg224 (łańcuch B) ustawienie łańcucha bocznego w taki sposób, by był on skierowany w stronę miejsca aktywnego. Takie ułożenie łańcucha bocznego Arg224 umożliwia solidne kotwiczenie substratu/produktu w miejscu aktywnym. Grupa guanidynowa Arg224 kotwiczy L-Asp za pomocą dwóch

wiązań wodorowych w mostku solnym pomiędzy atomami $N_{\eta 1}$ i $N_{\eta 2}$ a grupą α-karboksylową liganda. Położenie łańcuchów bocznych His117 i Arg224 jest stabilizowane przez trzecią resztę przełącznika, Glu250 (z łańcucha B). Grupa β-karboksylowa Glu250 związana jest wodorowo z grupą guanidynową Arg224, tworząc z nią dwa równoległe wiązania wodorowe w mostku solnym $O_{\varepsilon 1}\cdots N_{\varepsilon}$ i $O_{\varepsilon 2}\cdots N_{\eta 2}$, oraz z pierścieniem aromatycznym His117 wiązaniem $O^{\varepsilon 1}\cdots N$. Dla porządku i czytelności dalszych rozważań należy tu przypomnieć, że jon potasu w pętli aktywującej (łańcuch C) jest koordynowany przez 4 reszty pętli oraz dwie cząsteczki wody utrzymywane w odpowiedniej pozycji przez łańcuch boczny ostatniej reszty pętli, Ser118, tworzącej z nimi wiązania wodorowe o długościach 3.21 i 3.68 Å.



Rysunek 7.13. Oddziaływania pomiędzy Ser118 i Glu250 w przełączniku AD struktury PvAspG1-K. Kolorami oznaczyłam łańcuchy polipeptydowe: A - jasnozielony i D - ciemnozielony, przy czym atomy N są niebieskie, O czerwone. Zielona kula oznacza jon potasu w pętli aktywującej, czerwone kule cząsteczki wody, czarne linie przerywane wiązania koordynacyjne jonu metalu w pętli aktywującej. Czerwone linie przerywane obrazują niektóre z oddziaływań w obszarze przełącznika katalitycznego AD modelu PvAspG1-K.

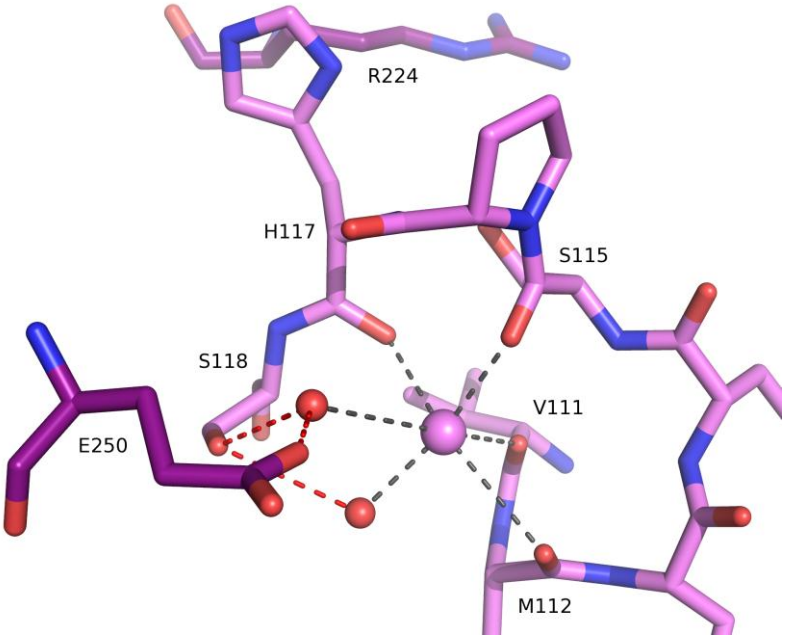
Drugi przełącznik modelu PvAspG1-K (przełącznik AD), kontaktujący pętlę aktywującą (łańcuch A), koordynującą jon potasu i miejsce aktywne (łańcuch D) nie obsadzone ligandem, jest również „włączony”. Reszty His117 (z łańcucha A) i Arg224 (z łańcucha D) są ułożone identyczne jak w przełączniku CB PvAspG1-K. Konformacja Glu250 (z łańcucha D) jest minimalnie zmieniona, ale jej łańcuch boczny zajmuje to samo miejsce w przestrzeni i nadal pełni funkcje stabilizującą dla dwóch pozostałych reszt przełącznika. Grupa β-karboksylanowa Glu250 (a konkretnie jej atom $O_{\varepsilon 1}$) oddziałuje z N_{ε} i $N_{\eta 2}$ Arg224 oraz z pierścieniem imidazolowym His117 (Rysunek 7.13). Wszystkie trzy reszty przełącznika AD mają takie ułożenie jak w przełączniku CB.

Pozwala ono na kotwiczenie L-Asp w miejscu aktywnym. Zatem cząsteczka liganda nie ma wpływu na ułożenie reszt przełącznika katalitycznego, którym steruje tylko i wyłącznie rodzaj jonu skoordynowanego w pętli aktywującej. Co ciekawe, grupa β-karboksylanowa Glu250 (atom $O_{\varepsilon 2}$) z przełącznika AD tworzy również dodatkowe oddziaływania z łańcuchem bocznym oraz grupą amidową reszty Ser118 (Rysunek 7.13). Jednocześnie wiązania pomiędzy atomem O_{γ} Ser118 i cząsteczkami wody koordynującymi jon potasu w pętli aktywującej słabną (odległości: 3.74 i 4.14 Å).

Model PvAspG1-Na, zawierający jony sodu w pętlach aktywujących, jest przykładem enzymu z wyłączoną (zredukowaną) aktywnością asparaginazową (stan OFF), ponieważ reszty tworzące przełączniki katalityczne CB i AD mają zmienione konformacje. Mechanizm działania przełącznika opiszę etapami, porównując PvAspG1-K (przełącznik CB w stanie ON) i PvAspG1-Na (przełącznik CB w stanie OFF) (Rysunek 7.12), przy czym należy pamiętać, że w rzeczywistości wszystkie opisywane zmiany następują najprawdopodobniej równocześnie, a nie sekwencyjnie. Gdy jon potasu zostaje zastąpiony jonem sodu, pętla aktywująca (np. w łańcuchu C) ulega rearanżacji. Zmienia się przebieg jej łańcucha głównego oraz następują zmiany konformacyjne kilku reszt budujących pętlę. Reszta His117 (łańcuch C), a właściwie jej pierścień imidazolowy zostaje niejako „wepchnięty” do wnętrza białka, zajmując miejsce do tej pory zajmowane przez łańcuch boczny Arg224 (łańcuch B). Dla uniknięcia zawady sterycznej, reszta Arg224 rotuje, zmieniając położenie łańcucha bocznego o 180º, który traci przez to możliwość kotwiczenia produktu/substratu w miejscu aktywnym. Jednocześnie następuje przemieszczenie znajdującej się w pobliżu, pętli Met218-Gly223 (tworzonej przez łańcuch B) oraz zmiana położenia łańcucha bocznego Met221. Trzecia reszta przełącznika katalitycznego, Glu250 (łańcuch B), także zmienia swoją konformację. β-Karboksylanowa grupa reszty Glu250 w nowej orientacji nie ma już możliwości oddziaływania z resztami aminokwasowymi białka (Rysunek 7.14). Łańcuch boczny tej reszty tworzy jedyne wiązanie wodorowe z cząsteczką wody, która jest unieruchomiona w przestrzeni przez łańcuch boczny Ser118 (łańcuch C) i koordynuje jon sodu w pętli aktywującej (łańcuch C).

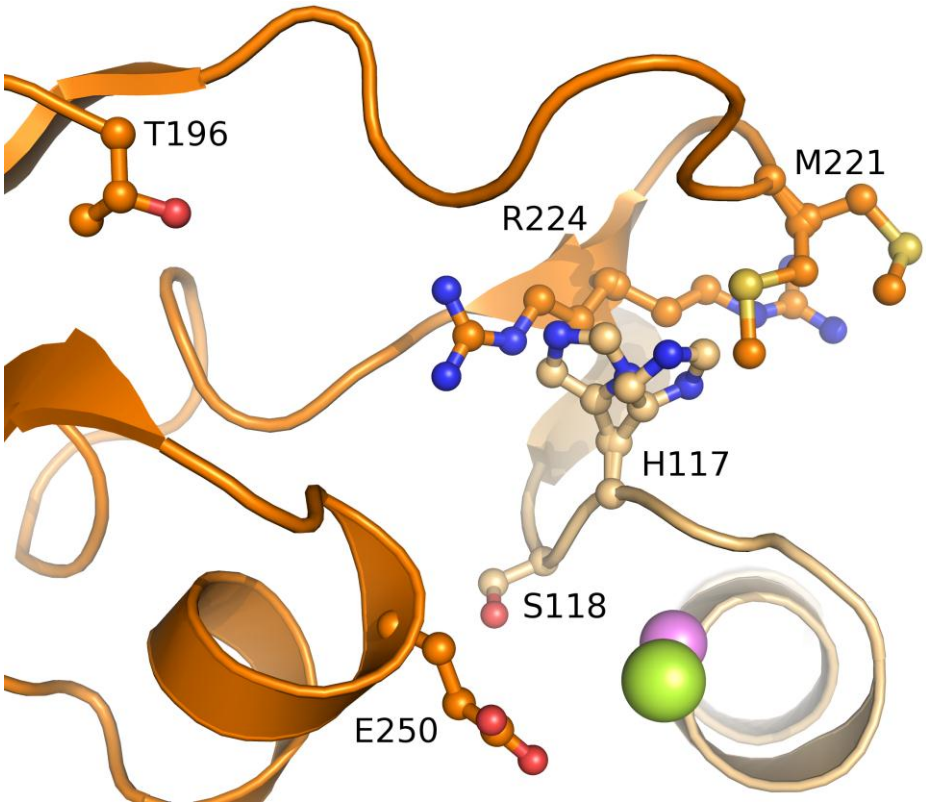
Przełącznik AD modelu PvAspG1-Na ukształtowany jest w identyczny sposób jak CB. Istnieją jedynie dwie drobne różnice, a mianowicie: (I) rearanżacja położenia cząsteczek wody znajdujących się pomiędzy jonem sodu w pętli aktywującej (łańcuch A) i resztą Glu250 (łańcuch D), oraz (II) fragmentaryczna mapa gęstości elektronowej

łańcucha bocznego Glu250, któremu nadałam konformację wyłączoną, podobną do tej w łańcuchu B.



Rysunek 7.14. Oddziaływania pomiędzy Ser118 i Glu250 w przełączniku CB struktury PvAspG1-Na. Kolorami oznaczyłam łańcuchy polipeptydowe: C - jasnofioletowy i B - ciemnofioletowy, przy czym atomy N są niebieskie, O czerwone. Fioletowa kula oznacza jon sodu, czerwone kule cząsteczki wody, czarne linie przerywane wiązania koordynacyjne jonu metalu w pętli aktywującej. Czerwone linie przerywane obrazują niektóre z wiązań wodorowych w obszarze przełącznika katalitycznego CB modelu PvAspG1-Na.

Rysunek 7.15. Zbliżenie obszaru przełącznika katalitycznego PvAspG1-K/Na. Kolorami oznaczyłam łańcuch C - jasnopomarańczowy i B - ciemnopomarańczowy, przy czym atomy N są niebieskie, O czerwone, S żółte. Jon potasu reprezentuje jasnozielona kula, a jon sodu jasnofioletowa kula.



W modelu PvAspG1-K/Na reszty przełączników katalitycznych przyjmują identyczne konformacje w obu heterodimerach struktury. W tym przypadku, ze względu na podwójne obsadzenie pętli aktywującej (jonami K^+ 30% i Na^+ 70%) mamy do czynienia z sytuacją pośrednią, a raczej ze współistnieniem, w tym samym krysztale, kopii białka w konformacji ON i w konformacji OFF (Rysunek 7.15). Łańcuchy boczne His117 i Arg224 posiadają podwójną konformację przyjmując położenia analogiczne do tych w modelach PvAspG1-K (z obsadzeniem 30%) i PvAspG1-Na (70%), a więc zarówno konformację włączoną jak i wyłączoną. Reszta Glu250 ma nową konformację, choć położenie jej łańcucha bocznego jest zasadniczo zbieżne z tym w PvAspG1-Na, bez oddziaływań z innymi resztami aminokwasowymi. Podobnie, konformacja pętli Met218-Gly223 jest taka jak w PvAspG1-Na, przy czym łańcuch boczny reszty Met221 ma podwójną konformację.

7.3. Dyskusja wyników

7.3.1. Analiza porównawcza różnych asparaginaz typu roślinnego

7.3.1.1. Porównanie sekwencji aminokwasowych

Asparaginazy typu roślinnego dzielone są na dwie podgrupy: enzymy potasowo-zależne i potasowo-niezależne, choć sekwencje aminokwasowe obu podgrup wykazują znaczne podobieństwo i identyczność wynoszące odpowiednio ok. 70% i ok. 60% (Tabela 7.13). Sekwencje enzymów danej podgrupy są jeszcze bardziej zgodne, podobieństwo wynosi nawet ok. 95%, identyczność dochodzi do ok. 90%. Wyjątkowa zbieżność sekwencji aminokwasowej w podgrupie enzymów K^+-zależnych istnieje pomiędzy asparaginazą z fasoli zwyczajnej (PvAspG1) i asparagiazą z *Lotus japonicus* (LjNSE1). Oba enzymy mają taką samą liczbę reszt aminokwasowych wynoszącą 326 i tylko 29 różnic w sekwencji, przy czym 15 z tych reszt określa się jako podobne. Homologi asparaginaz typu roślinnego u organizmów innych niż roślinne są dużo bardziej odmienne. Zarówno identyczność, jak i podobieństwo ich sekwencji, w porównaniu do enzymów z roślin, jest obniżone do ok. 40%.

Tabela. 7.13. Porównanie sekwencji aminokwasowych wybranych asparaginaz typu roślinnego (część niezacieniowana - % identyczności; część zacieniowana - % podobieństwa). Pogrubioną czcionką zaznaczyłam enzymy potasowo-zależne, niepogrubioną potasowo-niezależne, pochyłą homologi asparaginaz potasowo-niezależnych typu roślinnego nie pochodzące z roślin (skróty reprezentujące poszczególne enzymy zamieściłam w rozdziale 1. Wykaz stosowanych skrótów).

	PvAspG1	**LjNSE1**	**At3g16150**	LlA	PvAspGT2	LjNSE2	At5g08100	*EcAIII*	*hANSase3*	*gpASNase3*
PvAspG1	-	91	85	60	59	59	59	40	38	40
LjNSE1	95	-	83	70	59	57	58	57	51	41
At3g16150	90	89	-	72	59	73	57	57	53	41
LlA	71	58	58	-	81	83	74	46	41	42
PvAspGT2	71	70	72	90	-	89	77	46	41	41
LjNSE2	73	72	57	91	94	-	78	63	56	41
At5g08100	70	70	69	82	83	83	-	61	55	40
EcAIII	56	41	40	61	63	46	45	-	38	38
hANSase3	50	40	41	55	56	41	41	53	-	80
gpANSase3	51	53	53	55	55	55	55	55	90	-

7.3.1.2. Porównanie struktur czwartorzędowych

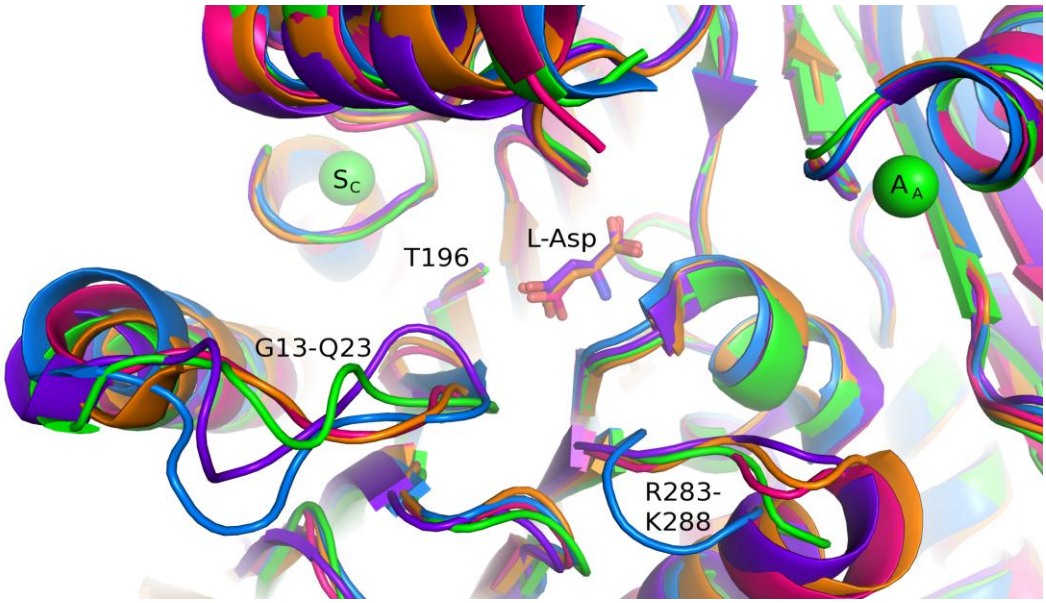
Porównanie struktur asparaginaz typu roślinnego (PvAspG1, LlA, EcAIII, hASNase3, gpASNase3) pokazuje, że mimo różnic w sekwencjach ich struktura czwartorzędowa jest bardzo podobna. Największa zgodność występuje między blisko spokrewnionymi enzymami, ssaczymi hASNase3 i gpASNase3, oraz roślinnymi PvAspG1 i LlA (Tabela 7.14).

Tabela. 7.14. Porównanie struktur heterotetramerów asparginaz typu roślinnego wykonane za pomocą programu ALIGN (**Satow *et al.*, 1986**). Podane wartości to r.m.s.d. (Å) dla atomów C_α. Porównałam pozycje atomów C_α struktur: PvAspG1-K (model enzymu zbadanego w niniejszej pracy, w stanie ON i ze związanym L-Asp w miejscu aktywnym, rozdzielczość 2.3 Å), LlA (kod PDB: 2gez - jedyny model aparaginazy z łubinu żółtego, rozdzielczość 1.9 Å), EcAIII (kod PDB: 2zal - model z *E. coli* o największej kompletności i ze związanym L-Asp w miejscu aktywnym, rozdzielczość 1.9 Å), hASNase3 (kod PDB: 4gdw - jedyny model w pełni dojrzałego enzymu ludzkiego ze związanym L-Asp w miejscu aktywnym, rozdzielczość 1.84 Å), gpASNase3 (kod PDB: 4o48 - jedyny model aparaginazy ze świnki morskiej, rozdzielczość 2.29 Å).

	LlA (2gez)	EcAIII (2zal)	hASNase3 (4gdw)	gpASNase3 (4o48)
PvAspG1-K	0.61	0.90	1.02	1.03
LlA (2gez)		0.82	0.96	1.02
EcAIII (2zal)			1.03	1.01
hASNase3 (4gdw)				0.46

Topologia heterodimerów αβ PvAspG1 jest identyczna z topologią LlA (**Michalska *et al.*, 2006**). Asparaginaza EcAIII zawiera dwie dodatkowe helisy: αH1 i βH4 (**Michalska *et al.*, 2005**). Natomiast enzymy hASNase3 i gpASNase3 posiadają dodatkową nić β w większym arkuszu β (βS10 - oznaczoną oryginalnie jako M; **Nomme *et al.*, 2012; Schalk i Lavie, 2014**). Co zaskakujące, w modelach ssaczych nie uwzględniono innej nici β: βS3 znajdującej się na przeciwnym skraju większego arkusza β, który spina dwa heterodimery αβ asparaginazy. Pominięcie βS3 było prawdopodobnie spowodowane niedoskonałością oddziaływania (obecne tylko dwa wiązania wodorowe -CO···HN-) między łańcuchami αS4 i βS3, co mogło utrudnić rozpoznanie tego ważnego elementu drugorzędowego przy pomocy automatycznych algorytmów.

Rysunek 7.16. Obszar asparaginaz typu roślinnego o najbardziej zróżnicowanym pofałdowaniu: pętla Gly13-Gln23 znajdująca się w podjednostce α i pętla Arg283-Lys288 z podjednostki β. Nałożyłam atomy C_α heterodimerów CD struktur: PvAspG1-K (zielony), LlA (niebieski), EcAIII (fioletowy), hASNase3 (pomarańczowy), gpASNase3 (różowy). Zielone kule reprezentują jony potasu w pętlach stabilizującej i aktywującej struktury PvAspG1-K, numeracja reszt aminokwasowych pochodzi także ze struktury PvAspG1-K. W miejscu aktywnym osadzony jest produkt reakcji (L-Asp) pochodzący ze struktur EcAIII, hASNase3, gpASNase3 (w kolorach łańcuchów białkowych).



Poza wspomnianymi powyżej różnicami, nałożenie wszystkich atomów C_α pięciu struktur asparaginaz ujawniło dodatkowo kilka mniej zgodnych fragmentów łańcucha białkowego, które wynikają głównie z insercji reszt aminokwasowych oraz naturalnej elastyczności pętli, czy helis znajdujących się przy powierzchni białek. Tylko dwa z tych obszarów można uznać za znacząco różniące się, ponieważ ich łańcuchy główne mają zupełnie inne pofałdowanie: pętlę Arg283-Lys288 w podjednostce β i pętlę Gly13-Gln23 w podjednostce α (numeracja wg PvAspG1) (Rysunek 7.16). Pętla Gly13-Gln23

znajduje się tuż za tzw. obszarem bogatym w Gly o wysoce zachowawczej strukturze. Konformacje tych pętli mogą być związane ze zróżnicowaną specyficznością substratową asparaginaz lub/i ich (nie)zależnością od potasu. Rejony te analizuję dokładniej w kolejnych rozdziałach.

W strukturach asparaginaz występuje jeszcze jeden fragment, co do którego istnieje podejrzenie, że jest wyznacznikiem preferencji substratowych (**Gabriel *et al.*, 2012**). Jest to tzw. pętla zmienna (*ang. variable loop*), C-terminalny fragment polipeptydu podjednostki α, który tworzy w niedojrzałym białku element łączący podjednostki heteodimerów i jest uwalniany w trakcie cięcia autoproteolitycznego. Pętla zmienna jest najbardziej zróżnicowanym fragmentem struktur asparaginaz pod względem sekwencji oraz ilości reszt aminokwasowych. Niestety większość modeli asparaginaz (również wszystkie trzy modele PvAspG1 prezentowane w niniejszej rozprawie) nie posiada tego fragmentu ze względu na jego nieuporządkowanie lub degradację. Jedynie w modelu LlA (**Michalska *et al.*, 2006**) w dwóch podjednostkach α (z czterech występujących w części asymetrycznej kryształu LlA) udało się wymodelować o kilka reszt aminokwasowych więcej, stanowiących ok. 25% pętli zmiennej. Nałożenie modeli PvAspG1, LlA, EcAIII, hASNase3 i gpASNase3 pokazuje, że dobudowany w LlA fragment jedynie okrywa miejsca aktywne i nie mógłby oddziaływać bezpośrednio z resztami tworzącymi miejsce aktywne, pętle koordynujące metale, czy przełącznik katalityczny w strukturze PvAspG1. Co ciekawe, łańcuch boczny reszty Thr165 znajdującej się w pętli zmiennej LlA tworzy słabe wiązanie wodorowe z resztą Gly10, która (jako Gly9 w hASNase3) pełni przypuszczalnie istotną role podczas dojrzewania białek hASNase3 i gpASNase3. Konformacja oraz potencjalne oddziaływania tworzone przez ten niezdefiniowany w modelach fragment pętli zmiennej pozostają nieznane.

7.3.1.3. Pętle koordynujące jony metali

W strukturach LlA, EcAIII, hASNase3 i gpASNase3 istnieje tylko jeden rodzaj pętli wiążących jon metalu - tzw. pętle sodowe (we wszystkich modelach tych struktur koordynujące jony sodu) odpowiadające pętli stabilizującej PvAspG1. Sekwencje aminokwasowe tych pętli różnią się (Tabela 7.15), jednak nie ma to znaczącego wpływu na ich funkcję ze względu na fakt, że koordynacja centralnie umieszczonego jonu metalu odbywa się wyłącznie za pomocą grup karbonylowych łańcucha głównego reszt

tworzących pętlę. Różne rodzaje grup bocznych decydują jedynie o dopasowaniu pętli do otaczających ją elementów strukturalnych. Przebieg łańcucha głównego oraz geometria sfery koordynacyjnej jonu są zatem podobne. Reszta Asn65, która w PvAspG1 odpowiada za stabilizujące oddziaływanie z katalityczną resztą Thr196 jest w innych strukturach asparaginaz zachowana, ma tą samą konformację i pozycję w przestrzeni. Pełni także tę samą funkcję we wszystkich w pełni dojrzałych (pociętych na podjednostki α i β) enzymach (choć nie we wszystkich heterodimerach modeli hASNase3). Pętla stabilizująca posiada zatem taką konformację, która umożliwia koordynowanie różnych jonów metali i (w przeważającej większości przypadków) zachowanie funkcji stabilizującej.

Reszty, które w PvAspG1 tworzą pętle aktywującą (Val111-Ser118), w strukturach innych asparaginaz typu roślinnego również tworzą podobne pętle, przy czym obsadzenie tych pętli jest zaskakująco zróżnicowane. W modelu LlA (który zawiera dwa heterotetramery asparaginazy w części asymetrycznej komórki elementarnej) pętle te są albo puste (dwie pętle) albo zawierają cząsteczki wody (dwie pętle) (**Michalska *et al.*, 2006**). W modelach EcAIII, których w bazie PDB jest sześć (12 podjednostek α), w pętle zostały wbudowane cząsteczki wody (osiem pętli), cząsteczka Tris (jedna pętla), jon Ca^{2+} (jedna pętla) oraz jony Na^{+} (dwie pętle modelu o kodzie PDB: 3c17). Obecność jonów sodu w pętlach aktywujących EcAIII (kod PDB: 3c17) to jednak artefakt, nie mający nic wspólnego z funkcją tej pętli w PvAspG1. Model ten zawiera jeszcze dziewięć jonów sodu związanych w różnych częściach struktury, w tym w obu miejscach aktywnych, w obu pętlach sodowych/stabilizujących i w pięciu dodatkowych pozycjach na powierzchni białka. Takie wysycenie struktury jonami sodu jest efektem bardzo dużego stężenia chlorku sodu (4.3 M) w roztworze krystalizacyjnym (**Michalska *et al.*, 2008**). W modelach hASNase3 (kody PDB: 4gdt, 4gdu, 4gdv, 4gdw, 4hlo, 4hlp, 4o0c, 4o0d, 4o0e, 4o0f, 4o0g, 4o0h) pętle aktywujące są puste (dwie pętle), zawierają cząsteczki wody (trzynaście pętli) lub cząsteczki glicyny (osiem pętli), której używano podczas krystalizacji. Jedna pętla hASNase3 (kod PDB: 4o0c) zawiera jon sodu, przy czym kryształy te przepłukiwano 75% nasyconym NaCl. Co ciekawe, w innych modelach hASNase3, których kryształy traktowano w ten sam sposób (**Nomme *et al.*, 2014**) są pętle z prawdopodobnie błędnie wymodelowanymi cząsteczkami wody zamiast jonów Na^{+} (co wynika z weryfikacji geometrii oraz długości wiązań). Modele gpASNase3 (PDB: 4o47, 4o48) zawierają w pętlach aktywujących cząsteczki wody (cztery pętle) (**Schalk i Lavie, 2014**). Zróżnicowane

obsadzenie pętli aktywujących asparaginaz potasowo-niezależnych jest oczywiście skutkiem procesu przygotowywania oraz krystalizacji białek. Ewidentne jest, że pętle aktywujące (w sprzyjających warunkach) mogą związać jony obecne w środowisku, jednak tylko w enzymie potasowo-zależnym PvAspG1 koordynacja jonów metali alkaicznych jest specyficzna, jako że wszystkie trzy modele asparaginazy z fasoli zwyczajnej niezależnie od warunków krystalizacji takie jony zawierają.

Tabela. 7.15. Analogiczne reszty aminokwasowe w różnych strukturach asparaginaz typu roślinnego. Reszty tworzące oktaedryczne miejsce koordynujące jony metali w pętlach są zaznaczone pogrubioną czcionką, reszty nie posiadające analogicznej funkcji są zaznaczone kursywą.

		PvAspG1	LlA	EcAIII	hASNase3	gpASNase3
miejsce aktywne	katalityczny nukleofil	Thr196	Thr193	Thr179	Thr168	Thr168
	triada treoninowa (Rozdział 7.3.2)	Thr196 Thr214 Thr246	Thr193 Thr211 Thr243	Thr179 Thr197 Thr230	Thr168 Thr186 Thr219	Thr168 Thr186 Thr219
	reszty budujące wnękę oksyanionową	Thr246 Gly247	Thr243 Gly244	Thr230 Gly231	Thr219 Gly220	Thr219 Gly220
	reszty wiążące substrat/produkt grupa α-aminowa	Asp227 Gly247	Asp224 Gly244	Asp210 Gly231	Asp199 Gly220	Asp199 Gly220
	reszty wiążące substrat/produkt grupa α-karboksylowa	Arg224 Gly249	Arg221 Gly246	Arg207 Gly233	Arg196 Gly222	Arg196 Gly222
	obszar bogaty w Gly (zawierający motyw GAG oraz pętle HGG)	His9 Gly10 Gly11 Ala12 Gly13	His9 Gly10 Gly11 Ala12 Gly13	His9 Gly10 Gly11 Ala12 Gly13	His8 Gly9 Gly10 Gly11 Ala12	His8 Gly9 Gly10 Gly11 Ala12
pętla stabilizująca (wokół pozycji $S_{A/C}$ jonów)		**Leu58** **Glu59** Thr60 **Asp61** Pro62 Leu63 **Phe64** Asn65 **Ser66** Gly67 **Arg68**	Leu59 **Glu60** Asn61 **Ile62** Glu63 His64 **Phe65** Asn66 **Ala67** Gly68 **Ile69**	**Leu60** **Glu61** Glu62 **Cys63** Pro64 Leu65 **Phe66** Asn67 **Ala68** Gly69 **Ile70**	**Leu55** **Glu56** Asp57 **Asp58** Pro59 Glu60 **Phe61** Asn62 **Ala63** Gly64 **Cys65**	**Leu55** **Glu56** Asp57 **Asp58** Ala59 Glu60 **Phe61** Asn62 **Ala63** Gly64 **His65**
pętla aktywująca (wokół pozycji $A_{A/C}$ jonów)		**Val111** **Met112** Asp113 Lys114 **Ser115** Pro116 **His117** Ser118	*Val112* *Met113* *Asp114* *Lys115* *Thr116* *Pro117* *His118* *Ile119*	*Val113* *Met114* *Glu115* *Gln116* *Ser117* *Pro118* *His119* *Val120*	*Val108* *Met109* *Glu110* *Lys111* *Thr112* *Pro113* *His114* *Cys115*	*Val108* *Met109* *Asp110* *Lys111* *Thr112* *Pro113* *His114* *Cys115*
przełącznik katalityczny		His117 Arg224 Glu250	*His118* *Arg221* *Glu247*	*His119* *Arg207* *Glu234*	*His114* *Arg196* *Glu223*	*His114* *Arg196* *Glu223*

Pomimo zróżnicowanego obsadzenia oraz niezależnie od różnic w sekwencjach (Tabela 7.15) pętle aktywujące w strukturach czterech asparaginaz potasowo-niezależnych (LlA, EcAIII, hASNase3, gpASNase3) przyjmują bardzo podobną konformację, przypominającą sytuację w strukturze asparaginazy z fasoli zwyczajnej, z grupami karbonylowymi skierowanymi do środka i łańcuchami bocznymi skierowanymi na zewnątrz. Co więcej, przebieg łańcucha głównego jest zbliżony do tego w PvAspG1-K, a kluczowa reszta pętli, His117, będąca jednocześnie elementem przełącznika katalitycznego jest zachowana i posiada konformację permanentnie włączoną (ON).

Porównując sekwencje różnych asparaginaz typu roślinnego (Rysunek 7.17) zauważyłam, że ostatnią resztą tworzącą pętlę aktywującą w enzymach potasowo-zależnych jest Ser, a w enzymach potasowo-niezależnych Ile, Val lub Cys. W PvAspG1 Ser118 oddziałuje z cząsteczkami wody koordynującymi jon metalu w pętli aktywującej oraz prawdopodobnie pośredniczy w zmianach konformacyjnych jednej z reszt przełącznika katalitycznego (Glu250). Wydaje się zatem, że zarówno specyficzne powinowactwo do metali alkaicznych, jak i możliwość włączania i wyłączania przełącznika katalitycznego jest związana z obecnością Ser118 w pętli aktywującej asparaginaz potasowo-zależnych.

Rysunek 7.17. Porównanie sekwencji aminokwasowych obszaru pętli aktywującej różnych asparaginaz typu roślinnego. Reszty tworzące pętle w PvAspG1 zaznaczyłam żółtym tłem; kolorem czerwonym zaznaczyłam His117, zielonym Ser118 (tak jak i analogiczne reszty w innych asparaginazach). Gwiazdki oznaczają reszty identyczne, kropki i dwukropki reszty odpowiednio mniej i bardziej podobne. Porównanie wykonałam w programie *ClustalO* (**McWilliam *et al.*, 2013**).

```
PvAspG1     100 TVKNPISLARL VMDKSPHS YIAFSGAEDFAR 130 | enzymy
LjNSE1      100 TVKNPISLARL VMDKSPHS YLAFDGAEDFAR 130 | K+-zależne
At3g16150   100 TVKNPISLARL VMDKSPHS YLAFSGAEDFAR 130 |

PvAspGT2    100 TVVNAISLARL VMEKTPHI YLAFDGAEEFAR 130 | enzymy
LjNSE2      100 TVVNAVSLARL VMEKTPHI YLAFDGAEEFAR 130 | K+-niezależne
At5g08100   101 TVVNPISLARL VMEKTPHI YLAFDAAEAFAR 131 |
LlA         101 TVLNPISLARL VMDKTPHI YLAFQGAQDFAK 131 |

EcAIII      102 HLRNPVLAARL VMEQSPHV MMIGEGAENFAF 132 | homologi
hASNase3     97 CIANPIKLARL VMEKTPHC FLTDQGAAQFAA 127 | nie pochodzące
gpASNase3    97 CIANPIKLARL VMDKTPHC FLTGQGAAKFAA 127 | z roślin
                  : * :  *** **:::*   :  ..*  **
```

7.3.1.4. Miejsca aktywne i przełącznik katalityczny

Porównanie miejsc aktywnych PvAspG1-K, LlA, EcAIII, hASNase3 i gpASNase3 pokazuje znaczną zbieżność sekwencyjną i strukturalną niezależnie od

obecności cząsteczki L-Asp. Co więcej, jeżeli w miejscu aktywnym znajduje się produkt reakcji, jak w EcAIII (kod PDB: 2zal), hASNase3 (kod PDB: 4gdw) i gpASNase3 (kod PDB: 4o48) ma on taką samą orientację i konformację jak produkt w strukturze PvAspG1-K. Miejsca aktywne PvAspG1-Na i PvAspG1-K/Na odróżniają się jednym ważnym szczegółem, tj. położeniem łańcucha bocznego reszty Arg224 przełącznika katalitycznego.

Trzy reszty aminokwasowe tworzące w modelu PvAspG1 przełącznik katalityczny (His117, Arg224, Glu250) również są zachowane w sekwencjach LlA, EcAIII, hASNase3 i gpASNase3 (Tabela 7.15). Ponadto, we wszystkich modelach asparaginaz potasowo-niezależnych reszty te przyjmują takie same konformacje jak reszty przełącznika katalitycznego CB modelu PvAspG1-K (ON). Można zatem powiedzieć, że enzymy potasowo-niezależne (a przynajmniej poznane struktury LlA, EcAIII, hASNase3 i gpASNase3) są permanentnie włączone, niezależnie od obsadzenia pętli aktywujących.

7.3.2. Autokatalityczne dojrzewanie białka

Wszystkie asparaginazy typu roślinnego podlegają dojrzewaniu, które jest procesem autokatalitycznym i polega na rozcięciu prekursorowego łańcucha polipeptydowego na podjednostki α i β. Dojrzewanie jest procesem niezbędnym dla aktywności enzymatycznej asparaginaz, ponieważ dzięki cięciu uwalniana jest katalityczna reszta treoniny, która staje się pierwszą resztą na N-końcu podjednostki β. Proces dojrzewania zachodzi jednak dość różnie. Rekombinowane białko EcAIII dojrzewa w kilka godzin, LlA potrzebuje kilku dni (**Michalska i Jaskólski, 2006**), a hASNase3 i gpASNase3 wymagają jeszcze więcej czasu. Preparaty białek ssaczych w zasadzie zawsze pozostają niedotrawione, chyba że do preparatu białka zostanie dodany katalizator jakim jest glicyna (**Nomme *et al.*, 2012; Su *et al.*, 2013; Schalk i Lavie, 2014**).

Istnieją dwie koncepcje wyjaśniające mechanizm autokatalitycznego cięcia asparaginaz. Oba proponowane mechanizmy postulują, że grupa hydroksylowa łańcucha bocznego treoniny (która po cięciu staje się N-końcową resztą podjednostki β i jednocześnie nukleofilem reakcji katalitycznych przeprowadzanych przez te enzymy) atakuje własne N-końcowe wiązanie amidowe. Pierwsza koncepcja została zaproponowana na podstawie struktury krystalicznej niedojrzałego białka EcAIII z

mutacją w miejscu aktywnym Thr179Ala (**Michalska *et al.*, 2008**). Mechanizm ten wskazuje niezbędne dla procesu elementy strukturalne: obecność katalitycznej reszty Thr179 (w dojrzałym białku będącej resztą nukleofilową reakcji enzymatycznej), Thr230 pełniącej rolę zasady i Asn67 (z pętli sodowej / stabilizującej), której łańcuch boczny miałby być częścią wnęki oksyanionowej. Alternatywny mechanizm został zaproponowany na podstawie niedojrzałej (nieprzeciętej) struktury hASNase3, w której katalizator glicynowy, położony w bezpośrednim sąsiedztwie katalitycznej Thr168, pełni rolę zasady i inicjuje przecięcie łańcucha polipeptydowego (**Nomme *et al.*, 2012; Su *et al.*, 2013**). W dojrzewaniu ssaczych asparaginaz typu roślinnego biorą też udział dwie treoniny, które razem z katalityczną resztą (Thr168) tworzą tzw. triadę treoninową (**Nomme *et al.*, 2014; Schalk i Lavie, 2014**). Łańcuch boczny Thr186 jest donorem wiązania wodorowego dla Thr168, co ma sprawiać że proton grupy hydroksylowej katalitycznej reszty staje się bardziej labilny. Thr219 natomiast ma wpływać na położenie pętli HGG (znajdującej się w tzw. obszarze bogatym w Gly), w której konformacja glicyny (Gly9) zmienia się w zależności od tego, czy prekursorowy polipeptyd jest już przecięty, czy też nie.

Równie ważne dla procesu autokatalitycznego cięcia powinno być umiejscowienie pętli zmiennej, a przynajmniej tego fragmentu jej łańcucha, który tworzą reszty położone najbliżej katalitycznej treoniny. Jednak porównanie różnych modeli niedojrzałych asparaginaz typu roślinnego (kody PDB: 3c17, 2zak, 4gdu, 4hlo, 4o0c, 4o0d, 4o0e, 4o0f, 4o0g, 4o48) pokazuje dużą różnorodność konformacji tego fragmentu, prowadzącą do przeróżnych oddziaływań z katalityczną treoniną.

Jako że nie istnieje struktura niedojrzałej PvAspG1, w tym przypadku można tylko spekulować na temat specyficznego mechanizmu autokatalitycznej proteolizy. Jednak na podstawie informacji uzyskanych podczas oczyszczania białka można przypuszczać, że PvAspG1 dojrzewa najprawdopodobniej według mechanizmu zaproponowanego dla EcAIII. Próbka rekombinowanego białka PvAspG1 pobrana już na etapie izolacji z komórek gospodarza *E. coli* była formą dojrzałą. Obraz rozdziału elektroforetycznego (SDS-PAGE) pokazał dwa prążki o masie 21 kDa (podjednostka α) i 13.6 kDa (podjednostka β). Świadczy to o bardzo szybko postępującym procesie dojrzewania, prawdopodobnie zaraz po translacji - jeszcze w komórkach gospodarza. Co więcej, PvAspG1 nie potrzebuje glicyny do zajścia autoproteolizy. Katalizator glicynowy, niezbędny do autoproteolizy hASNase3 i gpASNase3 (również otrzymanych metodami biotechnologicznymi; **Nomme *et al.*, 2012, 2014; Su *et al.*,**

2013; Schalk i Lavie, 2014), prawdopodobnie nie jest (wystarczająco) dostępny w komórkach gospodarza (*E. coli*), jako że musiał być dodawany podczas oczyszczania w celu przyśpieszenia dojrzewania tych ssaczych homologów asparaginazy roślinnej.

7.3.3. Mechanizm reakcji katalitycznej

Asparaginazy typu roślinnego posiadają dwa rodzaje aktywności, przy czym substratem reakcji jest zawsze L-asparagina lub jej pochodna. Jako asparaginazy enzymy te katalizują reakcję hydrolizy L-asparaginy z uwolnieniem cząsteczki amoniaku. Jako izoaspartylopeptydazy katalizują reakcję hydrolizy wiązania β-amidowego pomiędzy dwoma aminokwasami β-peptydu.

Najnowszy mechanizm reakcji katalitycznej został zaproponowany przez **Nomme *et al.*** (**2014**) na podstawie analizy struktury krystalicznej hASNase3 (kod PDB: 4o0h), w której produkt reakcji (asparaginian) nadal pozostaje kowalencyjnie związany z katalityczną treoniną. Podstawowym elementem tego mechanizmu odpowiedzialnym za aktywność asparaginazową jest triada treoninowa (Thr168, Thr219, Thr186), co udowodniono za pomocą badań strukturalnych oraz biochemicznych z wykorzystaniem mutantów hASNase3 (**Nomme *et al.*, 2014**).

Według mechanizmu zaproponowanego przez **Nomme *et al.*, (2014**), reakcja rozpoczyna się nukleofilowym atakiem aktywowanego łańcucha bocznego Thr168 (N-końcowej reszty podjednostki β) na grupę karbonylową łańcucha bocznego substratu. Nukleofilowość Thr168 zapewniają dwa oddziaływania: łańcucha bocznego Thr186 będącego donorem wiązania wodorowego z grupą hydroksylową Thr168, oraz wolnej grupy α-aminowej samej katalitycznej treoniny, która pełni rolę zasady. Podczas ataku nukleofilowego na atomie tlenu łańcucha bocznego substratu w stanie przejściowym powstaje ładunek ujemny. Jest on stabilizowany przez wnękę oksyanionową, którą tworzą Thr219 i Gly220. Po uwolnieniu cząsteczki amoniaku, drugi produkt reakcji (asparaginian) pozostaje kowalencyjnie związany z Thr168 wiązaniem estrowym. Następnie aktywowana cząsteczka wody atakuje estrowy atom węgla i poprzez stan przejściowy uwalnia produkt.

Wnęka katalityczna w poznanych dotąd strukturach dojrzałych (pociętych na podjednostki α i β) asparaginaz z różnych organizmów (PvAspG1-K, LlA, EcAIII, hASNase3, gpASNase3) jest w zasadzie identyczna w budowie (zachowane są wszystkie reszty tworzące oddziaływania oraz ich konformacja) niezależnie od

obecności produktu. L-Asp osadzony w miejscu aktywnym przyjmuje taką samą konformację i tworzy takie same wiązania. Można zatem spodziewać się tego samego mechanizmu reakcji u wszystkich tych enzymów. Weryfikując mechanizm odnoszę się zatem nie tylko do prezentowanych w niniejszej rozprawie struktur PvAspG1, ale do wszystkich struktur asparaginaz typu roślinnego, zawężając jednak rozważania do tych modeli, które prezentują w pełni dojrzałe enzymy, ponieważ tylko takie asparaginazy są katalitycznie aktywne (kody PDB: 4pu6, 4pv2, 4pv3, 2gez, 1t3m, 1k2x, 2zal, 1jn9, 4gdw, 4hlp, 4o0h, 4o48).

Zaproponowany dla hASNase3 mechanizm pozostawia wątpliwości dotyczące aktywacji nukleofila (**Michalska i Jaskólski, 2006**). Po pierwsze wydaje się, że utworzenie wewnątrzcząsteczkowego wiązania wodorowego pomiędzy grupą hydroksylową i α-aminową katalitycznej treoniny jest mało prawdopodobne ze względu na niekorzystną geometrię takiego wiązania. Ponadto wolna grupa $-NH_2$ jest słabą zasadą w porównaniu do grupy -OH i łatwo ulega protonacji do formy $-NH_3$. Druga wątpliwość dotyczy wiązania wodorowego pomiędzy łańcuchami bocznymi Thr168 i Thr186 (numeracja wg hASNase3). Grupa hydroksylowa Thr186 jest akceptorem protonu w wiązaniach wodorowych z grupami -NH dwóch reszt Gly187 i Gly188 oraz donorem protonu w wiązaniu z Thr168. Katalityczna treonina przyjmuje zatem proton, a więc jej nukleofilowość zmniejsza się zamiast wzrastać. W mechanizmie tym prawdopodobnie należałoby uwzględnić więcej oddziaływań z otoczenia miejsca aktywnego, aby lepiej wytłumaczyć wzrost nukleofilowości katalitycznej treoniny.

Michalska *et al.* (**2005**) sugerowała na przykładzie EcAIII, że w oderwaniu protonu z katalitycznej treoniny mogłaby pomóc reszta asparaginy (Asn62 wg hASNase3) znajdująca się w pętli sodowej/stabilizującej, która w proponowanym mechanizmie zupełnie nie została wzięta pod uwagę. Stabilizujące wiązanie wodorowe pomiędzy wolną grupą α-aminową katalitycznej treoniny i karbonylem łańcucha głównego tej asparaginy istnieje we wszystkich rozważanych modelach asparaginaz. Niewykluczone jest też, że to wiązanie zapewnia odpowiednią do ataku nukleofilowego orientację przestrzenną katalitycznej treoniny. Również łańcuch boczny omawianej asparaginy mógłby brać udział w reakcji, ponieważ znajduje się on blisko grupy α-aminowej katalitycznej treoniny, jednak odległości między tymi grupami są zróżnicowane i nie zawsze stanowią o silnym wiązaniu wodorowym (3.0 - 4.2 Å). Alternatywnie, jak zaproponowali **Nomme *et al.*** (**2012**), łańcuch boczny Asn mógłby oddziaływać z cząsteczkami wody, które pośredniczyłyby w aktywacji treoniny przez

jej własną grupę α-aminową. Jednak zidentyfikowanie w różnych modelach asparaginaz charakterystycznych, katalitycznych, identycznie usytuowanych cząsteczek wody nie jest możliwe. Co ciekawe, analizowana reszta asparaginy z pętli stabilizującej położona jest również w pobliżu kolejnej reszty triady treoninowej (Thr186 wg hASNase3), która miałaby zapewniać nukleofilowość katalitycznej treoniny 168. Odległość od łańcucha bocznego Thr186 do łańcucha bocznego Asn62 wynosi 3.5 - 4.0 Å, a pomiędzy łańcuchem bocznym Thr186, a grupą karbonylową Asn62 3.2 - 4.1 Å. Ponadto istotną rolę funkcjonalną przewiduje się też dla usytuowanej obok miejsca aktywnego (w tzw. obszarze bogatym w Gly) pętli HGG (**Nomme *et al.*, 2014**). Grupa karbonylowa środkowej glicyny tej pętli zazwyczaj oddziałuje z grupą karbonylową łańcucha głównego katalitycznej treoniny. Interakcja ta w przypadku modelu hASNase3, w którym produkt związany jest kowalencyjnie, (kod PDB: 4o0h) jest zakłócona z powodu podwójnej konformacji Gly9, co ewidentnie sugeruje jej udział w reakcji. Tak więc, pomimo określenia licznych struktur przestrzennych asparaginaz typu roślinnego, jak dotąd nie ma w pełni przekonującego i powszechnie akceptowanego mechanizmu reakcji katalitycznej. Niezwykle pomocna byłaby możliwość analizy wysokorozdzielczej struktury asparaginazy z tej rodziny.

7.3.4. Specyficzność substratowa

Testy biochemiczne (przeprowadzone w Kanadzie przez grupę dr Marsolais) potwierdziły zależność maksymalnej aktywności asparaginazowej enzymu PvAspG1 od obecności jonów potasu (**Bejger *et al.*, 2014**). Nie wiadomo jednak, jakie są preferencje substratowe asparaginazy z fasoli zwyczajnej. Na podstawie wysokiego podobieństwa sekwencyjnego (95%) do enzymu z *Lotus japonicus* (LjNSE1) można przewidywać, że nie będzie ona hydrolizowała β-peptydów (**Credali *et al.*, 2011**). Co więcej, dzięki obecności L-Asp w jednym z miejsc aktywnych PvAspG1-K można także, poprzez porównywanie z innymi Ntn-hydrolazami, spekulować na temat potencjalnych substratów i określić cechy strukturalne decydujące o właściwościach biochemicznych tego enzymu.

Badania biochemiczne, jakie przeprowadzono dla białek LlA i EcAIII pokazały, że enzymy te nie mogą hydrolizować substratów, które posiadają modyfikacje grup α-karboksylowej lub α-aminowej (**Borek *et al.*, 2004**). Obserwacja ta zgodna jest z cechami strukturalnymi, ponieważ w PvAspG1-K oraz w innych asparaginazach typu

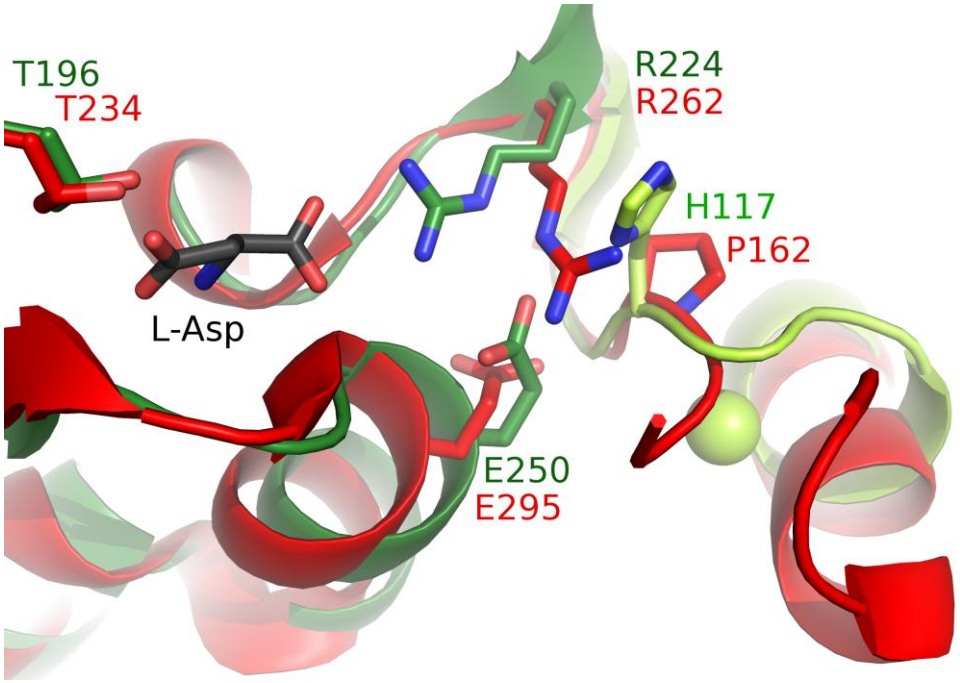
roślinnego wymienione grupy odpowiadają za kotwiczenie produktu (a więc i substratu) w miejscu aktywnym (Rysunek 7.7). Grupa α-aminowa oddziałuje z resztami Asp227 i Gly247, a grupa α-karboksylowa z Arg224 i Gly249. Dwa wiązania wodorowe jakie istnieją między grupą karboksylową L-Asp i grupą guanidynową Arg224 uniemożliwiają obracanie się produktu/substratu wokół własnej osi i precyzyjnie unieruchamiają go we wnęce katalitycznej. Odległość od końca łańcucha bocznego Arg224 do grupy hydroksylowej katalitycznej Thr196 definiuje dokładną długość, jaką powinien mieć substrat, aby hydrolizowane wiązanie znalazło się w odpowiednim dla ataku nukleofilowego miejscu (**Michalska i Jaskólski, 2006**). Tak więc, każda mała cząsteczka będąca pochodną L-asparaginianu, jak L-asparagina czy β-peptydy (np. β-Asp-Leu) może być potencjalnym substratem, jednak łańcuch boczny L-glutaminy jest już za długi, aby ulec hydrolizie przez LlA czy EcAIII (**Borek *et al.*, 2004**) i prawdopodobnie także przez PvAspG1-K.

Co ciekawe, nie wszystkie powyższe spostrzeżenia można odnieść do modelu PvAspG1-Na. Reszty przełącznika katalitycznego asparaginazy z fasoli zwyczajnej, która w pętli aktywującej koordynuje jon sodu, mają konformację wyłączoną. W tej sytuacji, łańcuch boczny Arg224 nie jest już skierowany w stronę miejsca aktywnego, przez co nie jest możliwe utworzenie wiązań pomiędzy Arg224 i (nieobecnym w tym modelu) produktem/substratem. Zakładając, że kontrolowana przez metale alkaiczne przemiana konformacyjna Arg224 jest cechą wszystkich asparaginaz potasowo-zależnych można przypuszczać, że to właśnie ona jest przyczyną zmiennej (a więc regulowanej) aktywności asparaginazowej tych enzymów. Pozostałe reszty biorące udział w wiązaniu L-Asp we wnęce katalitycznej nie zmieniają swojego położenia, stąd przy braku jonów K^+ enzymy te choć „osłabione" nadal są aktywne.

Do równie interesujących wniosków doszłam porównując PvAspG1 z taspazą (Tas1), która jest proteazą (aspartazą treoninową) katalizującą potranslacyjne dojrzewanie białek jądrowych MLL (*ang. mixed-lineage leukaemia*) poprzez cięcie łańcucha głównego za resztą asparaginianu (Asp-Gly). W analogicznym do obszaru przełącznika katalitycznego rejonie struktury taspazy pętla aktywująca w ogóle nie istnieje, dwie z trzech reszt przełącznika katalitycznego są zachowane (Arg i Glu), a trzecia (His) jest zmieniona na dużo mniejszą prolinę (**Khan *et al.*, 2005**). Łańcuch boczny argininy jest odchylony o ok. 60º od położenia skierowanego w stronę miejsca aktywnego (standardowego położenia „aktywnego" w asparaginazach) (Rysunek 7.18). Taka orientacja reszt aminokwasowych pozwala na związanie w miejscu aktywnym

Tas1 znacznie większego substratu, nawet fragmentu łańcucha głównego białka (konkretnie takiego z motywem -Xxx-(α-Asp)-Gly-Xxx-, jaki występuje w białkach jądrowych MLL). Możliwe więc, że zmiany konformacyjne reszt przełącznika katalitycznego wywołane zamianą jonów metali w pętlach, nie tylko regulują poziom aktywności asparaginazowej PvAspG1, ale także zmieniają preferencje substratowe tego enzymu. Jednak pomysł ten to tylko spekulacja, która powinna być zweryfikowana za pomocą badań biochemicznych oraz strukturalnych.

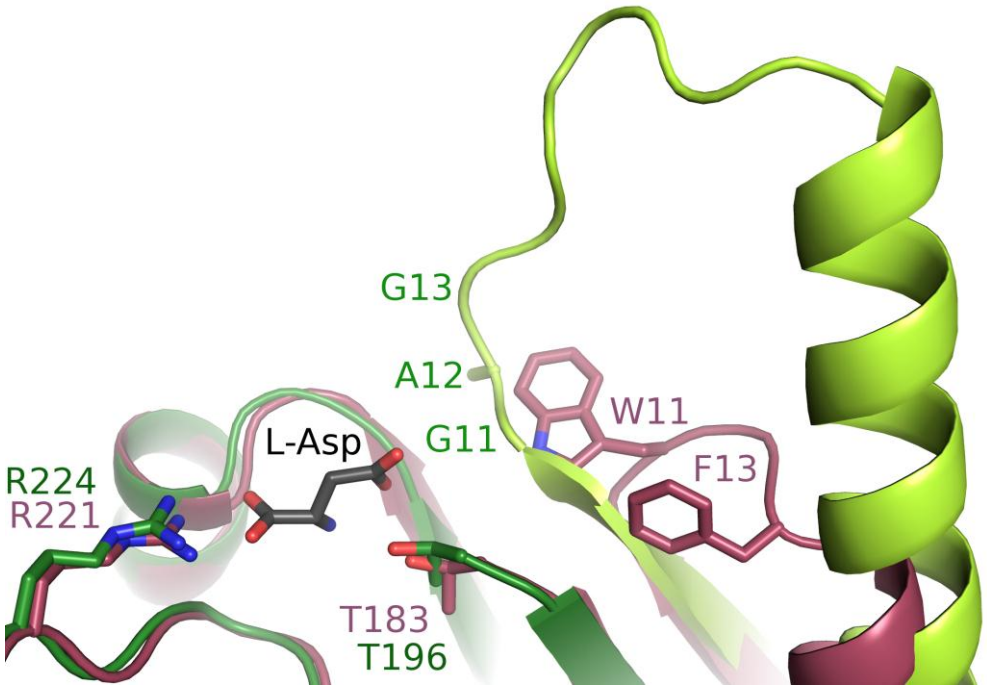
Rysunek 7.18. Porównanie rejonów obejmujących miejsce aktywne, przełącznik katalityczny i pętlę aktywującą struktury taspazy, Tas1 (kod PDB: 2a8j) (czerwona) ze strukturą PvAspG1-K (łańcuch C - jasnozielony, B - ciemnozielony, jon potasu - zielona kula, L-Asp w miejscu aktywnym – czarny kolor). Nałożyłam atomy C_α, kolorami zaznaczyłam atomy N - niebieskie, O - czerwone. Oznaczenia reszt aminokwasowych są w kolorach zgodnych z kolorystyką struktur.



Miejsce aktywne aspartyloglukozaminidazy (AGA), która jest kolejną Ntn-hydrolazą strukturalnie podobną do asparaginaz typu roślinnego, jest tak skonstruowane, aby mogło pomieścić N-β-glikozylowaną asparaginę. Wnęka katalityczna od strony łańcucha bocznego produktu zbudowana jest z krótkiej pętli zawierającej reszty aromatyczne (Trp11 i Phe13), które mają za zadanie wiązać grupę cukrową substratu (**Oinonen *et al.*, 1995; Guo *et al.*, 1998; Xuan *et al.*, 1998**). W PvAspG1 (jak i w innych asparaginazach typu roślinnego) obszar ten wypełnia długa pętla, dlatego tak duży substrat nie mógłby zmieścić się w jej miejscu aktywnym (Rysunek 7.19). Pętla ta

zawiera reszty Gly13-Gln23 (numeracja wg PvAspG1) i jest jednym z najbardziej zróżnicowanych (strukturalnie i sekwencyjnie) rejonów tych enzymów (Rysunek 7.16).

Rysunek 7.19. Porównanie rejonu miejsca aktywnego PvAspG1-K oraz AGA. Nałożyłam atomy C_α tych struktur, AGA jest przedstawiona kolorem różowym, PvAspG1-K kolorem zielonym (jasny - łańcuch A, ciemny - B) z L-Asp w miejscu aktywnym (czarny). Kolory oznaczeń reszt aminokwasowych są zgodne z kolorami struktur, atomy N są niebieskie, O czerwone.



Co ciekawe, rejon Gly13-Gln23 następuje po tzw. rejonie bogatym w Gly, w którym znajduje się pętla HGG oraz motyw GAG (Tabela 7.15). Pętla HGG, a dokładniej obie jej glicyny, w strukturach asparaginaz oddziałuje z grupą karbonylową katalitycznej treoniny. Według propozycji przedstawionej dla asparaginaz ssaczych (hASNase3 i gpASNase3), gra ona potencjalnie istotną rolę funkcjonalną zarówno podczas reakcji enzymatycznej jak i autokatalitycznego dojrzewania tych enzymów (**Nomme *et al.*, 2014; Schalk i Lavie, 2014**). Motyw GAG (nie występujący w gpASNase3, a w hASNase3 przesunięty o jedną resztę) ma prawdopodobnie oddziaływać z odcinaną resztą aminokwasową β-peptydowych substratów (**Michalska *et al.*, 2006**). Jako że cztery potasowo-niezależne asparaginazy (LlA, EcAIII, hASNase3, gpASNase3) potrafią hydrolizować oba substraty (L-Asn i β-peptydy, choć z różnym powinowactwem substratowym), a w stosunku do PvAspG1 istnieje podejrzenie (oparte na podobieństwie do LjNSE1), że enzym ten nie hydrolizuje β-peptydów, można by spodziewać się znacznych różnic między tymi białkami w obszarze bogatym w glicyny. Jednak cały ten rejon (mimo różnic sekwencyjnych) ma wysoce zachowawczą strukturę,

nie ma zatem oczywistej korelacji konformacji tego fragmentu łańcucha białkowego z dostępnymi informacjami na temat aktywności asparaginazowej i izoaspartylopeptydazowej.

Znajdująca się na C-końcu podjednostek α asparaginaz typu roślinnego pętla zmienna, również jest uznawana za obszar decydujący o preferencjach substratowych. Wymiana sekwencji tego rejonu pomiędzy enzymami potasowo-zależnym i niezależnym z *Arabidopsis thaliana* dokonana metodami inżynierii genetycznej znacząco wpłynęła na ich aktywność (**Gabriel *et al.*, 2012**). Niestety praktycznie we wszystkich strukturach krystalicznych dojrzałych asparaginaz obszar ten nie jest wymodelowany ze względu na brak dobrej jakości gęstości elektronowej. Tylko w modelu LlA widoczny jest fragment pętli okrywający miejsce aktywne (**Michalska *et al.*, 2006**). Brak możliwości porównania pętli zmiennych z różnych enzymów uniemożliwia weryfikację tezy o wpływie tego regionu na specyficzność substratową asparaginaz typu roślinnego.

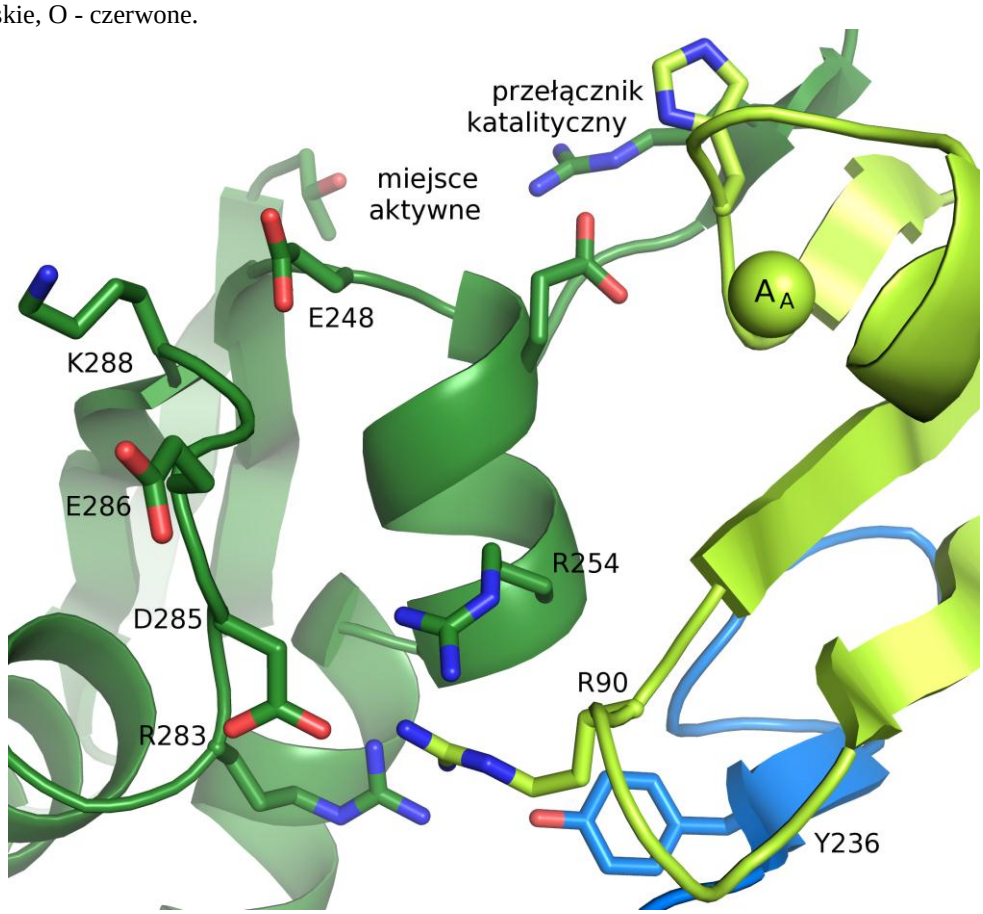
7.3.5. Struktura krystaliczna PvAspG1 na tle hipotez literaturowych

Przedstawiona w niniejszej rozprawie struktura PvAspG1 jest pierwszą strukturą potasowo-zależnej asparaginazy typu roślinnego zdefiniowaną na podstawie dyfrakcji promieni X na kryształach. Wcześniejsze doniesienia literaturowe mówiące o tego typu enzymach dotyczyły głównie badań kinetycznych. Istnieje także jedna publikacja (**Credali *et al.*, 2011**), w której autorzy wykorzystując narzędzia bioinformatyczne skonstruowali *in silico* model potasowo-zależnej asparaginazy z *Lotus japonicus* (LjNSE1) i na jego podstawie wskazali trzy reszty (Glu248, Asp285, Glu286) potencjalnie koordynujące jon K^+. Komputerowe modelowanie molekularne zostało uzupełnione ukierunkowaną mutagenezą trzech wskazanych reszt, które zostały zamienione na odpowiednie reszty (Lys248, Pro285, Lys286) występujące w sekwencji potasowo-niezależnej asparaginazy z *Lotus japonicus* (LjNSE2). Badania biochemiczne mutantów potwierdziły istotną rolę tych reszt dla aktywności enzymatycznej LjNSE1. Wykazano również, że aktywność asparaginazowa enzymu posiadającego wszystkie trzy mutacje przestaje zależeć od obecności jonów potasu. Niestety wygenerowana *in silico* struktura LjNSE1 nie jest dostępna w żadnej bazie struktur, stąd niemożliwe jest dokładne jej przeanalizowanie. Jednak w oparciu o ewolucyjne pokrewieństwo i niezwykłą zgodność sekwencyjną enzymów LjNSE1 i PvAspG1 (95% podobieństwa,

taka sama liczba i, co za tym idzie, numeracja poszczególnych reszt aminokwasowych) można założyć równie wysoką zgodność struktur i skonfrontować doniesienia dotyczące bioinformatycznego modelu ze strukturą ustaloną eksperymentalnie.

W PvAspG1 trzy wskazane reszty aminokwasowe Glu248, Asp285 i Glu286 są zachowane i umiejscowione obok siebie w obszarze sąsiadującym z miejscem aktywnym (Rysunek 7.20). Dwie z nich znajdują się w pętli Arg283-Lys288, która jest jednym z dwóch najbardziej zmiennych strukturalnie obszarów asparaginaz typu roślinnego (Rysunek 7.16). Trzecia, Glu248, mieści się w pętli pomiędzy nicią βS5 (zawierającą Thr246 i Gly247, które oddziałują z L-Asp) i helisą βH1 (zawierającą Gly249, która wiąże L-Asp, oraz Glu250, czyli element przełącznika katalitycznego).

Rysunek 7.20. Obszar zawierający Glu248, Asp285, Glu286 w PvAspG1-K. Łańcuchy: D - ciemnozielony, A - jasnozielony, B - jasnoniebieski, potas w pętli aktywującej - zielona kula, atomy N - niebieskie, O - czerwone.



W trzech przedstawionych przeze mnie modelach PvAspG1 wszystkie sześć reszt Glu248 ma dokładnie taką samą konformację zdefiniowaną przez jednoznaczne mapy gęstości elektronowej. Reszta Glu248 nie oddziałuje z resztami tworzącymi miejsce aktywne, czy z produktem reakcji. Nie uczestniczy też, tak jak to sugerowali **Credali *et***

al., (**2011**) w koordynacji jonu potasu. Obecne w PvAspG1 jony metali znajdują się w odległości ok. 13 Å (z pętli aktywujących) i ok. 20 Å (z pętli stabilizujących) od atomu C_α Glu248. Łańcuch boczny reszty Glu248 oddziałuje z grupą amidową łańcucha głównego Lys288. W strukturze potasowo-niezależnej asparaginazy z łubinu żółtego (LlA) w analogicznej pozycji znajduje się reszta Lys245, której łańcuch boczny nie tworzy żadnych znaczących wiązań wodorowych (**Michalska *et al.*, 2006**). Mutacja Glu248 na Lys248 w LjNSE1 powodowała znaczące (pięćdziesięciokrotne) obniżenie aktywności asparaginazowej enzymu (**Credali *et al.*, 2011**). Taka sama mutacja w PvAspG1 nie powinna powodować znaczących zmian w strukturze.

Reszty Asp285 i Glu286 są elementami elastycznej pętli Arg283-Lys288, której łańcuch główny ma taki sam przebieg we wszystkich trzech modelach, we wszystkich αβ heterodimerach PvAspG1. W pętli tej brakuje jednak gęstości elektronowej dla wielu łańcuchów bocznych. Niemożliwe jest zatem dokładne porównanie konformacji reszt aminokwasowych pętli Arg283-Lys288 w różnych modelach PvAspG1. Z tej przyczyny, przeanalizowałam tylko konformacje Asp285 i Glu286 na podstawie reszt, które posiadają kompletne mapy gęstości elektronowej.

Łańcuch boczny Asp285 jest w pełni widoczny w obu heterodimerach modelu PvAspG1-K. Oddziałuje on z łańcuchami bocznymi reszt Arg254 i Arg90 (Rysunek 7.20). Odległość do najbliższego jonu K^+ (w pętli aktywującej) wynosi ok. 16 Å. Łańcuch boczny Glu286 jest stuprocentowo zdefiniowany tylko w łańcuchu D modelu PvAspG1-Na. Oddziałuje on z łańcuchem bocznym reszty Arg254, a odległość do najbliższego jonu Na^+ (w pętli aktywującej) wynosi ok. 13 Å. Porównanie Asp285 i Glu286 z analogicznymi resztami innych asparaginaz typu roślinnego (LlA, EcAIII, hASNase3, gpASNase3) nie jest możliwe ze względu na zbyt duże różnice w przebiegu łańcuchów głównych pętli Arg283-Lys288 w tych enzymach. Mutacje Asp285Pro i Glu286Lys (**Credali *et al.*, 2011**) w PvAspG1 prawdopodobnie spowodowałyby destabilizację pętli Arg283-Lys288, ponieważ zanikłyby oddziaływania z Arg254 i Arg90.

Reasumując, Glu248, Asp285 i Glu286 w PvAspG1 nie oddziałują z resztami tworzącymi miejsce aktywne, przełącznik katalityczny, czy pętle wiążące jony metali (Rysunek 7.20). W ich bezpośrednim otoczeniu nie ma gęstości elektronowej świadczącej o obecności jonów metali, a ich liniowe ułożenie (Rysunek 7.20) na powierzchni struktury uniemożliwia stworzenie sfery koordynacyjnej o odpowiedniej geometrii. Co ciekawe, w strukturach białek jony sodu i potasu cztery razy częściej

koordynowane są przez grupy karbonylowe łańcucha głównego niż przez łańcuchy boczne Glu i Asp (**Harding *et al.*, 2010**). Stąd różnice w sekwencjach asparaginaz niekoniecznie muszą wskazywać reszty aminokwasowe biorące udział w koordynacji jonów metali alkaicznych.

W pętli Arg283-Lys288, oprócz reszt Asp285 i Glu286 występuje jeszcze jedna ważna reszta: Arg283. Biorąc pod uwagę wszystkie trzy modele PvAspG1, łańcuch boczny tej reszty posiada doskonałą gęstość elektronową w czterech z sześciu przypadków. Arg283 jest jedyną resztą struktury, która oddziałuje aż z trzema łańcuchami białka jednocześnie. Przykładowo, Arg283 z łańcucha D oddziałuje z Gly255 i Thr256 z tego samego łańcucha (D), z Arg90 z łańcucha A i Tyr236 z łańcucha B (Rysunek 7.4). Jeśli Arg283 (w dwóch kopiach: B i D) stanowi centrum spójności heterotetrameru PvAspG1, to możliwe jest, że wszelka destabilizacja pętli Arg283-Lys288 (np. mutacje Glu248Lys, Asp285Pro lub/i Glu286Lys) będzie wpływać na strukturę czwartorzędową PvAspG1 i w konsekwencji na aktywność enzymatyczną białka.

8. Podsumowanie

„Otrzymywanie i analiza krystalograficzna wybranych białek układu roślina motylkowata – rizobium" jest pracą doktorską, w której przedstawiam wyniki prac nad trzema niezależnymi zadaniami składowymi. Poszczególne zadania dotyczyły różnych obiektów badań, tj. różnych białek układu symbiotycznego i obejmowały inny zakres prac stanowiących fragment wieloetapowego procesu określania struktury przestrzennej białek. Stąd, w każdym z zadań pojawiały się inne problemy do rozwiązania.

8.1. Bakteryjne białka nodulacji

W pierwszym z zadań, obiektem badań była grupa pięciu białek nodulacji: NodA, NodB, NodU, NolO i NoeI, pochodzących z symbiotycznych bakterii brodawkowych. Celem zadania było przetestowanie wybranych białek pod kątem ich przydatności do badań krystalograficznych, a więc podjęcie próby uzyskania kryształów tych białek, bez których zastosowanie rentgenografii strukturalnej nie jest w ogóle możliwe. Szczegółowy zakres prac obejmował: uzyskanie plazmidowych wektorów ekspresyjnych, produkcję białek w komórkach *E. coli*, ich oczyszczanie i krystalizację.

Głównym problemem w tym zadaniu była nierozpuszczalność obiektów badań. Starałam się pokonać tę trudność na różne sposoby:

- wykorzystując odmienność sekwencji białek pochodzących z różnych szczepów;
- stosując wektory plazmidowe nadające różne cechy produkowanym białkom;
- zmieniając fizyko-chemiczno-biologiczne warunki procesu ekspresji (m. in. temperaturę, skład pożywek, komórki gospodarza);
- podejmując próbę ponownego zwinięcia białek izolowanych z ciałek inkluzyjnych.

Spośród wszystkich przetestowanych rozwiązań dwa okazały się kluczowe: termiczna indukcja białek opiekuńczych następująca przed ekspresją w 18ºC oraz obecność kowalencyjnie związanego białka fuzyjnego MBP. Dzięki tym dwóm zabiegom trzy białka, NodA, NodB i NoeI, mogły zostać poddane procesowi oczyszczania. Niestety, mimo moich usilnych starań, dwa pozostałe obiekty badań (NodU i NolO) pozostawały nierozpuszczalne uniemożliwiając mi przejście do następnego etapu badań.

Kolejnym problemem w tym zadaniu była niestabilność białek nodulacji, które bardzo łatwo wytrącały się z roztworów. Starałam się ograniczyć agregację modyfikując

protokoły oczyszczania, dodając lub usuwając poszczególne etapy. Modyfikowałam także skład stosowanych roztworów: używając buforów o różnym pH, optymalizując stężenie soli, wprowadzając czynniki redukujące oraz substancje poprawiające rozpuszczalność białek (np. glicerol, detergenty). Jedynie w przypadku białka NodA związanego z białkiem fuzyjnym MBP udało mi się skutecznie opanować agregację.

Aby zachować rozpuszczalność i stabilność białko NodA musiało pozostać związane z MBP, co źle wpływało na homogenność preparatu (weryfikowaną przy użyciu pomiarów DLS) i zmniejszało szanse pojawienia się kryształów. Zmieniłam więc plazmidowy wektor ekspresyjny tak, aby rekombinowane połączone białka NodA i MBP tworzyły sztywną konstrukcję. Ponadto zaprojektowałam dedykowany dla MBP zestaw roztworów do krystalizacji. Pomimo starań, nie uzyskałam żadnych monokryształów białek nodulacji.

Nie wiadomo dlaczego system ekspresyjny *E. coli*, który spośród wszystkich dostępnych systemów jest najbliższy genetycznie rizobiom i teoretycznie powinien być najwłaściwszy do produkcji ich białek, w zasadzie nie pozwalał uzyskać rozpuszczalnych białek. Nie wiadomo również dlaczego białka te są tak niestabilne w roztworach. Zawarta w literaturze charakterystyka białek nodulacji jest zbyt skąpa, aby moc wyciągać rzetelne wnioski. To wszystko sprawia, że białka nodulacji są niezwykle trudnymi obiektami badań i większość z nich nie nadaje się do badań krystalograficznych, z jedynym wyjątkiem białka NodA, które daje pewne nadzieje na pozytywny wynik ewentualnie kontynuowanych prac.

8.2. Asparaginaza z *Rhizobium* etli (AnsA)

W drugim zadaniu obiektem badań był enzym pochodzący z bakterii glebowej *Rhizobium etli*, która jest symbiotycznym partnerem fasoli zwyczajnej, zaliczany do słabo poznanej rodziny asparaginaz typu *Rhizobium etli* (AnsA). Celem zadania było umożliwienie podjęcia próby rozwiązania struktury tego enzymu, a więc przeprowadzenie obiektu przez te wszystkie etapy całego procesu określania struktury przestrzennej białek, które dotyczą pracy z materiałem białkowym: od ekspresji, poprzez oczyszczanie i krystalizację białka, aż do rejestracji danych dyfrakcyjnych.

Prace nad tym zadaniem rozpoczęłam przejmując wektor ekspresyjny pET-Re-AnsA umożliwiający ekspresję rekombinowanego białka AnsA w komórkach gospodarza *E. coli*. Produkcja rozpuszczalnego białka i opracowanie wydajnego

protokołu oczyszczania nie przysporzyły mi większych problemów. Uzyskiwane preparaty AnsA były dobrej jakości, homogenne i stabilne, co potwierdzały wyniki pomiarów DLS. Z kolei znalezienie i optymalizacja warunków krystalizacyjnych były zadaniami niezwykle pracochłonnymi oraz czasochłonnymi. Jednak wytrwale kontynuowałam próby mające na celu zmianę pokroju kryształów. Ostatecznie zamiast pozrastanych igieł i płytek uzyskiwałam osobne, duże płytki nadające się do pomiarów dyfrakcji. Najlepszy wyhodowany przeze mnie kryształ AnsA rozpraszał promieniowanie rentgenowskie do rozdzielczości 1.5 Å, a kluczowym odczynnikiem pozytywnie wpływającym na pokrój kryształów była oktyloglukoza.

W moich badaniach przyjęłam strategię wykorzystującą sygnał anomalny do rozwiązywania struktury krystalicznej. Niestety ingerencja w skład kryształów, polegająca na wprowadzeniu atomów specjalnych (nasączanie kompleksem tantalu, krystalizacja białka znakowanego selenometioniną) drastycznie pogarszała ich zdolność rozpraszania promieniowania rentgenowskiego. Podjęłam również próbę rejestracji słabego sygnału anomalnego pochodzącego od naturalnie występujących w białku atomów siarki, korzystając z najlepszych jakie udało mi się uzyskać natywnych kryształów AnsA. Niestety kryształy te cechowała anizotropia dyfrakcji, która znacząco pogarszała jakość danych dyfrakcyjnych. Z tego powodu zarejestrowanie statystycznie istotnego sygnału anomalnego nie było możliwe.

Korzystając z natywnego, rozpraszającego do ~1.5 Å kryształu białka AnsA zarejestrowano również jeden zbiór danych używając promieniowania o standardowej długości fali (0.91841 Å). Zbiór ten może posłużyć w przyszłości do rozwiązania struktury metodą podstawienia cząsteczkowego. W związku z przewidywaniami że, struktura AnsA znacznie różni się od struktury zbadanych dotąd asparaginaz, do metody MR najprawdopodobniej wykorzystany zostanie odpowiedni model wspomagany informacjami bioinformatycznymi.

Projekt dotyczący asparaginazy z *Rhizobium etli* jeszcze się nie zakończył i obecnie uczestniczą w nim kolejne osoby, wykonujące kolejne etapy prac niezbędne w procesie poznania struktury tego wyjątkowego enzymu.

8.3. Potasowo-zależna asparaginaza z *Phaseolus vulgaris* (PvAspG1)

W trzecim zadaniu obiektem badań była asparaginaza z *Phaseolus vulgaris* (PvAspG1), potasowo-zależny enzym typu roślinnego. Biochemiczna charakterystyka

białek z tej rodziny pokazuje, że obecne w środowisku jony metali alkaicznych wpływają na ich aktywność enzymatyczną: jony K^+ zwiększa ją, natomiast jony Na^+ zmniejsza ją. Celem zadania było wyjaśnienie podstaw molekularnych tej zależności. Szczegółowy zakres prac obejmował: rozwiązanie struktury przy użyciu metody podstawienia cząsteczkowego oraz udokładnienie i przeanalizowanie kilku modeli białka.

Prace rozpoczęłam wychodząc z zarejestrowanych na synchrotronie danych dyfrakcyjnych. W sumie, rozwiązane i udokładnione zostały aż trzy struktury tego samego białka, różniące się zawartością jonów: PvAspG1-K z jonami potasu, PvAspG1-Na z jonami sodu i PvAspG1-K/Na z oboma rodzajami jonów. Istnienie trzech modeli białka, które reprezentowały enzymy uchwycone w różnych stanach aktywności enzymatycznej umożliwiło wykonanie analizy porównawczej i zaowocowało ciekawymi wnioskami.

PvAspG1 ma zasadniczo taką samą strukturę jak inne asparaginazy typu-roślinnego, posiada jednak dwa rodzaje pętli, w których koordynowane są metale. Pętla stabilizująca (Leu58-Arg68), a właściwie występująca w jej sekwencji reszta asparaginy, „stabilizuje" położenie katalitycznej Thr196 i może być bezpośrednio lub pośrednio odpowiedzialna za mechanizm reakcji. Pętla ta, w enzymach potasowo-niezależnych, nazywana jest pętlą sodową, ale została przeze mnie przemianowana, ponieważ potrafi koordynować zarówno jony sodu jak i potasu nie tracąc swojej stabilizującej funkcji. Nie wiadomo jednak jaki wpływ miałyby inne jony jednowartościowe (np. Li^+ lub Rb^+) skoordynowane w tej pętli na jej strukturę i funkcję.

Pętla aktywująca (Val111-Ser118) jest specyficzną cechą strukturalną prezentowanego w niniejszej rozprawie enzymu potasowo-zależnego. We wcześniejszych strukturach potasowo-niezależnych asparaginaz pętla ta zazwyczaj nie wiąże jonów i nie pełni żadnej funkcji. Jon metalu koordynowany w pętli aktywującej PvAspG1 wpływa na aktywność enzymu za pośrednictwem mechanizmu opartego na zmianach konformacyjnych trzech reszt (His117, Arg224, Glu250) tworzących tzw. przełącznik katalityczny. Porównanie trzech modeli potasowo-zależnej PvAspG1 zawierających dwa różne rodzaje jonów (K^+ i Na^+) pokazuje różnice strukturalne jakie występują w dwóch stanach aktywności enzymu. Gdy pętla aktywująca koordynuje jon potasu (stan włączony) łańcuchy boczne reszt przełącznika przyjmują położenia umożliwiające stabilne kotwiczenie (przez Arg224) substratu/produktu w miejscu aktywnym. W strukturach enzymów potasowo-niezależnych taki układ przestrzenny reszt aminokwasowych występuje permanentnie. Gdy pętla aktywująca koordynuje jon

sodu (stan wyłączony) reszty przełącznika zmieniają konformacje powodując częściową destabilizację architektury miejsca aktywnego (od strony grupy α-karboksylowej substratu/produktu). Pozostałe elementy tworzące miejsce aktywne (np. katalityczna Thr196, reszty wnęki oksyanionowej, triada treoninowa, itp.) zachowują swoje położenia i pozostają poza wpływem koordynowanych jonów. Fakt ten wyjaśnia, dlaczego przy nieobecności jonów K^+ potasowo-zależne enzymy zmniejszają swoją aktywność asparaginazową, lecz nie tracą jej całkowicie. Wydaje się, że wymiana jonów K^+ na Na^+ w pętli aktywującej mogłaby również powodować zmianę preferencji substratowych PvAspG1, co zauważyłam porównując tą pętlę z analogicznym obszarem w Tas1, jednak hipoteza ta musi zostać zweryfikowana za pomocą badań biochemicznych.

Oba heterodimery αβ modelu PvAspG1-K mają pętle aktywujące obsadzone jonami potasu i oba przełączniki katalityczne w konformacji stanu włączonego, jednak tylko jedna wnęka katalityczna zawiera produkt reakcji L-Asp. Wnioskuję zatem, że obecność substratu/produktu w miejscu aktywnym nie wpływa na konformację reszt przełącznika i tylko jon koordynowany w pętli aktywującej steruje tym mechanizmem. Co więcej, ostatnia reszta budująca pętlę aktywującą, Ser118, jest najprawdopodobniej odpowiedzialna za istnienie tego mechanizmu zależnego od potasu. W sekwencjach asparaginaz potasowo-zależnych reszta ta jest zawsze zachowana, a w enzymach potasowo-niezależnych zawsze zmieniona. Łańcuch boczny Ser118 tworzy wiązania wodorowe z cząsteczkami wody koordynującymi jon metalu w pętli aktywującej oraz pośredniczy w zmianach konformacyjnych jednej z reszt przełącznika katalitycznego, Glu250. Postulat ten należy jednak zweryfikować przy pomocy testów biochemicznych oraz badań strukturalnych ukierunkowanych mutantów PvAspG1.

Analiza PvAspG1 wyjaśnia też znaczenie czwartorzędowej struktury tego enzymu. Potasowo-zależne asparaginazy mogą funkcjonować jedynie jako dimery heterodimerów $(\alpha\beta)_2$, ponieważ miejsce aktywne położone w jednym heterodimerze (np. AB) jest kontrolowane przez pętlę aktywującą z sąsiedniego heterodimeru (CD) za pośrednictwem przełącznika katalitycznego utworzonego z reszt pochodzących z obu tych podjednostek (His117C, Arg224B, Glu250B).

Rozwiązanie struktury krystalograficznej PvAspG1 wyjaśnia zagadkę zależności tego enzymu od potasu, ale także umożliwia postawienie kolejnych pytań dotyczących funkcjonowania asparaginaz. Czy ma znaczenie region bogaty w glicyny, zawierający motyw GAG potencjalnie odpowiedzialny za oddziaływanie z β-peptydami, skoro jego

sekwencja i budowa jest zachowana we wszystkich poznanych dotąd strukturach niezależnie od preferencji substratowych? Które cechy strukturalne odpowiedzialne są za rozpoznawanie różnych substratów? Jak pętla Arg283-Lys288 wpływa na aktywność enzymatyczną? Jak wzrasta nukleofilowość katalitycznej treoniny umożliwiająca przeprowadzenie reakcji? Czy możliwe jest zwiększenie aktywności tych enzymów i na ich bazie stworzenie nowych leków na białaczkę? Te wszystkie aspekty wymagają dalszych badań nad asparaginazami typu roślinnego.

9. Summary

"Preparation and crystallographic analysis of selected proteins from the legume plant-rhizobium symbiotic system" is a PhD thesis in which I have presented the results of my work on three independent but related projects. Each project concerned different research targets, i.e. different proteins from the symbiotic system studied, and each consisted of a different subset of the research tasks forming the multistage process of protein crystal structure determination, thus in each of them different problems had to be solved.

9.1. Bacterial nodulation proteins

In the first project, five nodulation proteins from symbiotic bacteria, including NodA, NodB, NodU, NolO and NoeI, were the research targets. The aim of this project was to test whether the selected proteins are suitable for crystallographic studies, or in other words it attempted to obtain single crystals adequate for X-ray diffraction experiments. The detailed scope of my work included: preparation of expression vectors, protein overexpression in *E. coli* bacterial cells, protein purification, and crystallization.

The main difficulty of this project was insolubility of the selected proteins. I've tried to overcome this barrier in several different ways:

- using sequence diversity of a given protein, derived from different species;
- utilizing plasmid vectors that introduced different features to the sequence of the recombinant proteins;
- changing the physico-chemical-biological conditions of the overexpression process, such as temperature, media composition, type (strain) of the host cells;
- trying to refold proteins isolated from inclusion bodies.

From all the tested solution, two turned out to be crucial, namely (i) thermal induction of chaperones before protein overexpression at 18ºC, and (ii) the presence of covalently bound MBP fusion protein. Thanks to these two strategies, three proteins: NodA, NodB and NoeI, could be successfully expressed and subjected to purification. Unfortunately, despite my efforts, the two remaining proteins (NodU and NolO) remained to be insoluble and in those two cases I have been unable to proceed to the next step of the

procedure.

Instability of the nodulation proteins, leading to precipitation, was the next problem with this task. I tried to reduce the aggregation by modifications of the purification protocols, and by adding or removing steps in the procedure. I was also modifying the composition of the buffers: using different pH, optimizing salt concentration, introducing reducing agents and factors that improve solubility, such as glycerol or detergents. Only in the case of the NodA protein bound to MBP, I was able to overcome the aggregation process.

To keep the NodA protein soluble and stable, it had to be fused with MPB, but this was affecting the sample homogeneity (as verified by the DLS measurements) and reducing the likelihood of crystallization. Therefore, I changed the expression vector to link the NodA and MBP proteins with a short, rigid element. Moreover, I designed a set of crystallization solutions dedicated to MBP. Despite these efforts, I was not able to get any single crystals of nodulation proteins.

It is not known, why the *E. coli* expression system, which among all the available systems is genetically the closest to the rhizobium organisms and theoretically should be most appropriate for the production of their proteins, essentially did not allow to obtain any soluble proteins. It is also unclear, why these proteins are so unstable in solutions. The characterization of nodulation proteins that can be found in the literature is too scanty to allow drawing any reliable conclusions. All this makes the nodulation proteins an extremely difficult research object. Most of them are not suitable for crystallographic studies, with the only possible exception of NodA, which gives some hope for a positive outcome if the work is continued.

9.2. Asparaginase from z *Rhizobium etli* (AnsA)

The second project concerned AnsA, an enzyme from *Rhizobium etli* - a soil bacterium which is a symbiont of common bean. The AnsA protein belongs to the poorly known subfamily of *Rhizobium etli*-type asparaginases. The aim of my work was to sound out and prepare ground for future work on the solution of the crystal structure of AnsA, or in other words, to carry out the project through all the stages that lead to crystallization and X-ray data collection.

I started the project by taking over the expression vector pET-Re-AnsA, which allows the overexpression of the AnsA protein in *E. coli* host cells. Production of

soluble protein and development of an efficient purification protocol was quite straightforward. The obtained AnsA preparations were of good quality, homogeneous and stable, as confirmed by DLS measurements. In contrast, finding and optimizing the crystallization conditions was very laborious and time-consuming. However, I persistently continued my attempts to improve the crystal morphology. Finally, instead of bundles of needles or plates, I managed to grow large, separated plates suitable for diffraction experiments. AnsA crystal of the best quality diffracted X-rays up to 1.5 Å. The critical reagent improving the crystals morphology was octyl-β-glucoside.

In my approach, I have adopted a strategy that would allow the use of anomalous signal to solve the crystal structure. Unfortunately, tempering with crystal composition, such as introduction of special atoms (by soaking with a tantalum complex, or crystallization of Se-Met labeled protein) has invariably had a deletorious effect on diffraction quality. I also made an attempt to register the weak anomalous signal from the naturally occurring sulfur atoms, by using my best native AnsA crystals. Unfortunately, those crystals showed a strongly anisotropic diffraction, effectively prohibiting recording of a high-quality data set. For these reasons, exploitation of the anomalous signal was not possible.

There is also one native data set, collected at a standard wavelength of 0.91841 Å to ~1.5 Å resolution. This data set can be used in the future to solve the crystal structure by bioinformatics-augmented molecular replacement. Sequence analyses predict that the AnsA structure is different from other asparaginase structures examined so far, thus a simple MR method is not likely to work.

I leave the AsnA project unfinished. Currently, other lab members will carry on, hoping that the structure of this unique enzyme will be cracked one day.

9.3. Potassium-dependent asparaginase from *Phaseolus vulgaris* (PvAspG1)

In the third project, the potassium-dependent plant-type asparaginase from *Phaseoulus vulgaris* (PvAspG1) has been studied. Biochemical characterization of proteins from this family shows that the alkali metal ions present in the environment influence their enzymatic activity: the K^+ ions increase the activity, while Na^+ decrease it. The goal of this project was to elucidate the molecular basics of this dependence. The detailed scope of my work included: structure solution by molecular replacement, refinement and analysis of several protein models.

I started my work by processing X-ray diffraction data collected at a synchrotron. In total, three crystal structures of the same enzyme were solved, differing in alkali metal ion content: (i) PvAspG1-K with potassium ions, (ii) PvAspG1-Na with sodium ions, and (iii) PvAspG1-K/Na with both ions. Three protein models that represent the enzyme caught in different states of the enzymatic process, allowed me to carry out a comparative analysis and resulted in interesting conclusions.

PvAspG1 has the same general fold as other plant-type asparaginases, but in variance with potassium independent enzymes has not one but two types of loops where metal ions are coordinated. The stabilization loop (Leu58-Arg68), or more precisely the Asn residue in its sequence, "stabilizes" the position of the catalytic Thr196 nucleophile and may be directly or indirectly responsible for the reaction mechanism. This loop in potassium-independent asparaginases is called "sodium loop", but I have renamed it as it is capable of coordinating both sodium as well as potassium ions without losing its stabilizing function.

The activation loop (Val111-Ser118) is a specific structural feature of the presented potassium-dependent enzyme. In previously reported structures of potassium-independent asparaginases this loop usually does not coordinate any ions and has no function. The metal ion coordinated in the activation loop of PvAspG1, affects the enzyme activity through a mechanism based on conformational changes of three key residues, His117, Arg224 and Glu250, forming the so-called catalytic switch. A comparison of the three models of the potassium-dependent PvAspG1 containing two different types of ions (K^+ or/and Na^+) shows the structural changes that occur in two activity states of the enzyme. When the activation loop coordinates a potassium ion in the ON state, the side chains of the catalytic switch residues adopt positions that enable productive anchoring (by Arg224) of the substrate/product in the active site. In potassium-independent enzymes such a spatial arrangement of the catalytic switch is permanent. When the activation loop coordinates a sodium ion (in the OFF state), the catalytic switch residues change their conformation leading to partial destabilization of the active site architecture at the α-carboxyl end of the substrate/product. Other elements forming the active site (e.g. the catalytic Thr196, the oxyanion hole residues, a threonine triad, etc.) maintain their positions and are not influenced by the activation-loop-coordinated metal ions. This explains why, when potassium ions are absent, the potassium-dependent enzymes slow down their asparaginase activity, but do not lose it completely. It seems that the K^+/Na^+ exchange in the activation loop could also

influence substrate preference of PvAspG1. However, this hypothesis has to be verified by biochemical studies.

Both αβ heterodimers of PvAspG1-K have their activation loops occupied with potassium ions and both catalytic switches are in the ON state, but only one active site pocket is occupied by a reaction product molecule, L-Asp. I, therefore, conclude that the presence of substrate/product in the active site does not affect the conformation of the switch residues, and that the switch state is only controlled by the potassium ion coordinated in the activation loop. Moreover, the last residue of the activation loop, Ser118, is most likely responsible for the existence of this potassium-dependent mechanism. Ser in this position is not present in potassium-independent asparaginases, but is strictly conserved in the potassium-dependent enzymes. The side chain of Ser118 forms hydrogen-bonds with water molecules coordinated by the metal ion in the activation loop and also mediates the conformational changes of one of the catalytic switch residue, Glu250. This postulate has to be verified by biochemical tests and structural studies of PvAspG1 mutants.

Analysis of PvAspG1 explains also the significance of its quaternary structure, i.e. why this potassium-dependent enzyme forms an obligatory $(\alpha\beta)_2$ dimer of heterodimers. Since the active site of one heterodimer (e.g. AB) is controlled by the activation loop from the other heterodimer (CD) through the catalytic switch residues Arg224 and Glu250 (from chain B) and His117 (from chain C), this mechanism can only work within a dimer of heterodimers.

The crystal structure of PvAspG1 explains the enzyme's potassium dependence mystery, but also allows to raise further questions concerning asparaginases in general. Does the glycine-rich region, containing a GAG motif potentially responsible for interaction with β-peptide substrates, really matters, as its sequence and structure are conserved in all known structures regardless of their of substrate preferences? Which structural features are responsible for the recognition of different substrates? How does the Arg283-Lys288 loop affect the enzymatic activity? How is the nucleophilicity of the catalytic threonine enhanced? Is it possible to increase the asparagine affinity of these enzymes and convert them to antileukemic drugs? All these aspects require further studies on plant-type asparaginases.

10. Bibliografia

1. Abdel-Lateif K., Bogusz D., Hocher V. (2012). The role of flavonoids in the establishment of plant roots endosymbioses with arbuscular mycorrhiza fungi, rhizobia and *Frankia* bacteria. *Plant Signal. Behav.* 7: 636-641.
2. Adams P.D., Afonine P.V., Bunkoczi G., Chen V.B., Davis I.W., Echols N., Headd J.J., Hung L.W., Kapral G.J., Grosse-Kunstleve R.W., McCoy A.J., Moriarty N.W., Oeffner R., Read R.J., Richardson D.C., Richardson J.S., Terwilliger T.C., Zwart P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* D66: 213-221.
3. Adxv (2013). Adxv - a program to display X-ray diffraction images. http://www.scripps.edu/~arvai/adxv.html.
4. Allen FH (2002). The Cambridge Structural Database: A quarter of a milion crystal structures and rising. *Acta Cryst.* B58: 380-388.
5. Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., LipmanD.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25: 3389-3402.
6. Amor B.B., Shaw S.L., Oldroyd G.E., Maillet F., Penmetsa R.V., Cook D., Long S.R., Denarie J., Gough. (2003). The NFP locus of *Medicago truncatula* controls an early step of nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J.* 34: 495-506.
7. Andrade A.F., Borges K.S., Silveira V.S. (2014). Update on the use of L-asparaginase in infants and adolescent patients with Acute Lymphoblastic Leukemia. *Clin. Med. Insights. Oncol.* 8: 95-100.
8. Ardourel M., Demont N., Debelle F., Maillet F., DeBilly F., Prome J.C., Denarie J., Truchet G. (1994). *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. *Plant Cell* 6: 1357-1374.
9. Arrighi J.F., Barre A., Ben Amor B., Bersoult A., Soriano L.C., Mirabella R., de Carvalho-Niebel F., Journet E.P., Gherardi M., Huguet T., Geurts R., Denarie J., Rouge P., Gough C. (2006). The *Medicago truncatula* lysin [corrected] motif-recepor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol.* 142: 265-279.
10. Arsene F., Tomoyasu T., Bukau B. (2000). The heat shock response of *Escherichia coli*. *Int. J. Food Microbiol.* 55: 3-9.
11. Atkinson E.M., Palcic M.M., Hindsgaul O., Long S.R. (1994). Biosynthesis of *Rhizobium meliloti* lipooligosaccharide nod factors: NodA is required for an *N*-acyltransferase activity. *Proc. Nat. Acad. Sci. USA* 91: 8418-8422.
12. Azevedo R.A., Lancien M., Lea P.J. (2006). The aspartic acid metabolic pathway,

an exciting and essential pathway in plants. *Amino Acids* 30: 143-162.

13. Bejger M., Imiołczyk B., Clavel D., Gilski M., Pajak A., Marsolais F., Jaskólski M. (2014). Na^+/K^+ exchange switches the catalytic apparatus of potassium-dependent plant L-asparaginase. *Acta Cryst*. D70: 1854-1872.

14. Bergfors T.M. (ed.). Protein Crystallization: techniques, strategies and tips: a laboratory manual. (1999) *Internat. Univ. Line, La Jolla USA*.

15. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissing H., Shindyalov I.N., Bourne P.E. (2000). The Protein Data Bank. *Nucleic Acid Res.* 28: 235-242.

16. Bloemberg G.V., Kamst E., Harteveld M., van der Drift K.M., Haverkamp J., Thomas-Oates J.E., Lugtenberg B.J.J., Spaink H.P. (1995). A central domain of *Rhizobium* NodE protein mediates host specificity by determining the hydrophobicity of fatty acyl moieties of nodulation factors. *Mol. Microbiol*. 16: 1123-1136.

17. Bondos S.E., Bicknell A. (2003). Detection and prevention of protein aggregation before, during and after purification. *Anal. Biochem*. 316: 223-231.

18. Bonthon D.T., Jaskólski M. (1997). Why a "benign" mutation kills enzyme activity. Structure-based analysis of the A176V mutant of *Saccharomyces cerevisiae* L-asparaginase I. *Acta Biochim. Pol.* 44: 491-504.

19. Borek D., Michalska K., Brzezinski K., Kisiel A., Podkowiński J., Bonthron D.T., Krowarsch D., Otlewski J., Jaskólski M. (2004). Expression, purification and catalytic activity of *Lupinus luteus* asparagine β-amidohydrolase and its *Escherichia coli* homolog. *Eur. J. Biochem*. 271: 3215-3226.

20. Borek D., Jaskólski M. (2001). Sequence analysis of enzymes with asparaginase activity. *Acta Biochim. Pol.* 48: 893-902.

21. Brannigan J.A., Dodson G., Duggleby H.J., Moody P.C., Smith J.L., Tomchick D.R., Murzin A.G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* 378: 416-419.

22. Brese, N. E., O'Keeffe, M. (1991). Bond-valence parameters for solids. *Acta Cryst.* B47, 192-197.

23. Broughton W.J., Jabbouri S., Perret X. **(**2000). Keys to symbiotic harmony. *J. Bacteriol.* 182: 5641-5652.

24. Brown D. (2009). Recent developments in the methods and applications of the bond valence model. *Chem. Rev*. 109: 6858-6919.

25. Bruneau L., Chapman R., Marsolais F. (2006). Co-occurrence of both L-asparaginase subtypes in *Arabidopsis:* At3g16150 encodes a K^+-dependent L-asparaginase. *Planta* 224: 668-679.

26. Brunger A.T. (1992). The free R value: a novel statistical quantity for assessing

the accuracy of crystal structures. *Nature* 355: 472-474.

27. Brzeziński K., Stępkowski T., Panjikar S., Bujacz G., Jaskólski M. (2007). High-resolution structure of NodZ fucosyltransferase involved in the biosynthesis of the nodulation factor. *Acta Biochim. Pol.* 54: 537-549.

28. Brzeziński, K. (2006). Badania biochemiczne i krystalograficzne SAHazy i fukozylotransferazy NodZ, białek z układu symbiotycznego roślina - bakteria. *Praca Doktorska, Wydział Chemii Uniwersytetu im. A. Mickiewicza w Poznaniu.*

29. Cakici O., Sikorski M., Stępkowski T., Bujacz G., Jaskólski M. (2010). Crystal structures of NodS *N*-methyltransferase from *Bradyrhizobium japonicum* in ligand-free form and as SAH complex. *J. Mol. Biol*. 404: 874-889.

30. Campbell H.A., Mashburn L.T., Boyse E.A., Old L.J. (1967). Two L-asparaginases from *Escherichia coli* B. Their separation, purification, and antitumor activity. *Biochem*. 6: 721-730.

31. Cantor J.R., Stone E.M., Chantranupong L., Georgiou G. (2009). The human asparaginase-like protein 1 hASRGL1 is an Ntn-hydrolase with beta-aspartyl peptidase activity. *Biochem.* 48: 11026-11031.

32. Chen V.B., Arendall III W.B, Headd J.J, Keedy D.A., Immormino R.M., Kapral G.J., Murray L.W., Richardson J.S., Richardson D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst*. D66: 12-21.

33. Cooper E.J. (2007). Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. *J. App. Microbiol*. 103: 1355-1365.

34. Credali A., Diaz-Quintana A., Garcia-Calderon M., De la Rosa M.A., Marquez A.J., Vega J.M. (2011). Structural analysis of K^+ dependence in L-asparaginases from *Lotus japonicus. Planta* 234: 109-122.

35. Dauter Z., Dauter M., Dodson E. (2002). Jolly SAD. *Acta Cryst*. D58: 494-506.

36. Debelle F., Plazanet C., Roche P., Pujol C., Savagnac A., Rosenberg C., Prome J.C., Denarie J. (1996). The NodA proteins of *Rhizobium meliloti* and *Rhizobium tropici* specify the *N*-acylation of nod factors by different fatty acids. *Mol. Microbiol.* 22: 303-314.

37. DeLano W.L. (2002). The PyMOL Molecular Graphics System. *DeLano Scientific, San Carlos, CA, USA.*

38. Demont N., Debelle F., Aurelle H., Denarie J., Prome J.C. (1993). Role of the *Rhizobium meliloti nodF* and *nodE* genes in the biosynthesis of lipooligosaccharidic nodulation factors. *J. Biol. Chem*. 268: 20134-20142.

39. Denarie J., Debelle F., Rosenberg C. (1992). Signalling and host range variation in nodulation. *Annnu. Rev. Microbiol*. 46: 497-531.

40. D'Haeze W., van Montagu M., Prome J.C., Holsters M. (1999). Carbamoylation of azorhizobial nod factors is mediated by NodU. *Mol. Plant Microbe Interact*.

12: 68-73.

41. D'Haeze W., Holsters M. **(**2002). Nod factor structures, responses and perception during initiation of nodule development. *Glycobiology* 12: 79-105.

42. Dodson G., Wlodawer A. (1998). Catalytic triads and their relatives. *Trends Biochem. Sci.* 23: 347-352.

43. Downie J.A. (2010). The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. *FEMS Microbiol. Rev*. 34: 150-170.

44. Dunlop P.C., Meyer G.M., Ban D., Roon R.J. (1978). Characterization of two forms of asparaginase in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 253: 1297-1304.

45. Egelhoff T.T., Long S.R. (1985). *Rhizobium meliloti* nodulation genes: identification of *nodDABC* gene products, purification of NodA protein, and expression of NodA in *Rhizobium meliloti*. *J. Bacteriol.* 164: 591-599.

46. Emsley P., Cowtan K. (2004). Coot: model-building tools for molecular graphics. *Acta Cryst.* D60: 2126-2132.

47. Endre G., Kereszt A., Kevei Z., Mihacea S., Kalo P., Kiss G.B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature* 417: 962-966.

48. Fine B.M., Kaspers G.J.L., Ho M., Loonen A.H., Boxer L.M. (2005). A genome-wide view of the in vitro response to L-asparaginase in acute lymphoblastic leukemia. *Cancer Res*. 65: 291-299.

49. Fox J.D., Kapust R.B., Waugh D.S. (2001). Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci.* 10: 622-630.

50. Fraysse N., Couderc F., Pointsot V. (2003). Surface polysaccharide involvement in establishing the rhizobium-legume symbiosis. *Eur. J. Biochem*. 270: 1365-1380.

51. Freiberg C., Fellay R., Bairoch A., Broughton W.J., Rosenthal A., Perret X. (1997). Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* 387:394-401.

52. Gabriel M., Telmer P.G., Marsolais F. (2012). Role of asparaginase variable loop at the carboxyl terminal of the alpha subunit in the determination of substrate preference in plants. *Planta* 235: 1013-1022.

53. Garman E. (1999). Cool data: quantity and quality. *Acta Cryst* D55: 1641-1653.

54. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. Walker J.M. (ed). The Proteomics Protocols Handbook, *Humana Press* 571-607.

55. Geelen D., Leyman B., Mergaert P., Klarskov K., van Montagu M., Geremia R.A.,

Holsters M. (1995). NodS is an *S*-adenosyl-L-methionine dependent methyltransferase that methylates chitooligosaccharides deacetylated at the non-reducing end. *Mol. Microbiol*. 17: 387-397.

56. Giraud E., Moulin L., Vallenet D., Barbe V., Cytryn E., Avarre J.C., Jaubert M., Simon D., Cartieaux F., Prin Y., Bena G., Hannibal L., Fardoux J., Kojadinovic M., Vuillet L., Lajus A., Cruveiller S., Rouy Z., Mangenot S., Segurens B., Dossat C., Franck W.L., Chang W.S., Saunders E., Bruce D., Richardson P., Normand P., Dreyfus B., Pignol D., Stacey G., Emerich D., Vermeglio A., Medigue C., Sadowsky M. Legumes symbioses: absence of *nod* genes in photosynthetic *Bradyrhrizobia*. *Science* 316: 1307-1312.

57. Gonzalez V., Santamaria R.I., Bustos P., Hernandez-Gonzalez I., Medrano-Soto A., Moreno-Hagelsieb G., Janga S.C., Ramirez M.A., Jimenez-Jacinto V., Collado-Vides J., Davila G. (2006). The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. *Proc. Nat. Acad. Sci. USA* 103: 3834-3839.

58. Gottfert M., Rothlisberger S., Kundig C., Beck C., Marty R., Hennecke H. (2001). Potential symbiosis-specific genes uncovered by sequencing a 410-kilobase DNA region of the *Bradyrhrizobium japonicum* chromosome. *J. Bacteriol.* 183: 1405-1412.

59. Gough C., Jacquet C. (2013). Nod factor perception protein carries weight in biotic interactions. *Trends in Plant Science* 18:566-574.

60. Graham P.H., Vance C.P. (2003). Legumes: importance and constraints to greater use. *Plant Physiol*. 131: 872-877.

61. Gresshoff P.M., Hayashi S., Biswas B., Mirzaei S., Indrasumunar A., Reid D., Samuel S., Toillenaere A., van Hameren B., Hastwell A., Scott P., Ferguson B.J. (2015). The value of biodiversity in legume symbiotic nitrogen fixation and nodulation for biofuel and food production. *J. Plant Physiol*. 172: 128-136.

62. Hamel L.P., Beaudoin N. (2010). Chitooligosaccharide sensing and downstream signaling: contrasted outcomes in pathogenic and beneficial plant-microbe interactions. *Planta* 232: 787-806.

63. Harding M.M. (2006). Small revisions to predicted distances around metal sites in proteins. *Acta Cryst*. D62: 678-682.

64. Harding M.M., Nowicki M.W., Walkinshaw M.D. (2010). Metals in protein structures: a review of their principal features. *Cryst. Rev*. 16: 247-302.

65. Hendrickson W.A., Ogata C.M. (1997). Phase determination from multiwavelength anomalous diffraction measurements. *Methods in Enzymology* 276: 494-523.

66. Huerta-Saquero A., Evangelista-Martinez Z., Moreno-Enriquez A., Perez-Rueda E. (2013). *Rhizobium etli* asparaginase II: an alternative for acute lymphoblastic leukemia (ALL) treatment. *Bioenginered* 4: 30-36.

67. Huerta-Zepeda A., Duran S., Du Point G., Calderon J. (1996). Asparagine degradation in *Rhizobium etli*. *Microbiology* 142: 1071-1076.

68. Huerta-Zepeda A., Ortuno L., Du Pont G., Duran S., Lloret A., Merchant-Larios H., Calderon J. (1997). Isolation and characterization of *Rhizobium etli* mutants altered in degradation of asparagine. *J. Bacteriol.* 179: 2068-2072.

69. Iantcheva A., Mysore K.S., Ratet P. (2013). Transformation of leguminous plants to study symbiotic interactions. *Int. J. Dev. Biol*. 57: 577-586.

70. Iiboshi Y., Papst P.J., Hunger S.P., Terada N. (1999). L-asparaginase inhibits the rapamycin-targeted signaling pathway. *Biochem. Biophys. Res. Commun.* 260: 534-539.

71. Indrasumunar A., Kereszt A., Searle I., Miyagi M., Li D., Nguyen C.D., Men A., Carroll B.J., Gesshoff P.M. (2010). Inactivation of duplicated nod factor receptor 5 (NFR5) genes in recessive loss-of-function non-nodulation mutants of allotetraploid soybean (*Glicine max* L. Merr). *Plant Cell Physiol.* 51: 201-214.

72. Indrasumunar A., Searle I., Lin M.H., Kereszt A., Men A., Carroll B.J., Gesshoff P.M. (2011). Nodulation factor receptor kinase 1α controls nodule organ number in soybean (*Glicine max* L. Merr). *Plant J.* 65: 39-50.

73. Jabbouri S., Fellay R., Talmont F., Kamalaprija P., Burger U., Relić B., Promé J.C., Broughton W.J. (1995). Involvement of *nodS* in N-metylation and *nodU* in 6-O-carbamoylation of *Rhizobium* sp. NGR234 nod factors. *J. Biol. Chem*. 270: 22968-22973.

74. Jabbouri S., Relić B., Hanin M., Kamalaprija P., Burger U., Promé D., Promé J.C., Broughton W.J. (1998). *nolO* and *noeI* (HsnIII) of *Rhizobium* sp. NGR234 are involved in 3-*O*-carbamoylation and 2-*O*-methylation of nod factors. *J. Biol. Chem.* 273: 12047-12055.

75. John M., Röhrig H., Schmidt J., Wieneke U., Schell J. (1993). *Rhizobium* NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. *Proc. Nat. Acad. Sci. USA* 90: 625-629.

76. Kabsh W. i Sander C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22: 2577-2637.

77. Kabsch W. (2010). XDS. *Acta Cryst*. D66: 125-132.

78. Kafetzopoulos D., Thireos G., Vournakis J.N., Bouriotis V. (1993). The primary structure of a fungal chitin deacetylase reveals the function for two bacterial gene products. *Proc. Nat. Acad. Sci. USA* 90: 8005-8008.

79. Kalo P., GleasonC., Edwards A., Marsh J., Mitra R.M., Hirsch S., Jakab J., Sims S., Long S.R., Rogers J., Kiss G.B., Downie J.A., Oldroyd G.E. (2005). Nodulation signalling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science* 308: 1786-1789.

80. Kamst E., Pilling J., Raamsdonk L.M., Lugtenberg B.J.J., Spaink H.P. (1997). *Rhizobium* nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in nod factor biosynthesis. *J. Bacteriol*. 179: 2103-2108.

81. Karamitros C.S., Konrad M. (2014). Human 60-kDa lysophospholipase contains an N-terminal L-asparaginase domain that is allosterically regulated by L-asparagine. *J. Biol. Chem*. 289: 12962-12975.

82. Kaźmierczyk M. Wstępne badania biochemiczne i krystalograficzne wybranych enzymów klasy hydrolaz. *Praca Magisterska, Wydział Chemii Uniwersytetu im. A. Mickiewicza w Poznaniu.*

83. Khan J.A., Dunn B.M., Tong L. (2005). Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Structure* 13: 1443-1452.

84. Kleywegt G.J., Jones T.A. (1997). Model building and refinement practice. *Methods in Enzymology* 277: 208-230.

85. Klock H.E., Lesley S.A. (2009). The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. Doyle S.A. (ed.), Methods in Molecular Biology: High Throughput Protein Expression and Purification, vol. 498, *Humana Press, Totowa, NJ*

86. Kondorosi E., Mergaert P., Kereszt A. (2013). A paradigm for endosymbiotic life: cell differentiation of *Rhizobium* bacteria provoked by host plant factors. *Annu. Rev. Microbiol.* 67: 611-628.

87. Kotzia G.A., Labrou N.E. (2007). L-asparaginase from *Erwinia chrysantemi* 3937:cloning, expression and characterization. *J. Biotechnol*. 127: 657-669.

88. Kouchi H., Imaizumi-Anraku H., Hayashi M., Hakoyama T., Nakagawa T., Umehara Y., Suganuma N., Kawaguchi M. (2010). How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. *Plant Cell Physiol.* 51: 1381-1397.

89. Kretovich V.L., Sidelnikova L.I., Ivanushkin A.G., Karayakina T.L. (1981). Localization of aspartase, asparaginase and glutaminase in intact bacteroids of *Rhizobium lupini*. *Prikl. Biokhim. Microbiol*. 20: 445-447.

90. Krissinel E., Henrick K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst*. D60: 2256-2268.

91. Kuppuraj G., Dudev M., Lim C. (2009). Factors governing metal-ligand distances and coordination geometries of metal complexes. *J. Phys. Chem*. B. 113: 2952-2960.

92. Landeta C., Davalos A., Cevallos M.A., Geiger O., Brom S., Romero D. (2011). Plasmids with a chromosome-like role in rhizobia. *J. Bacteriol*. 193: 1317-1326.

93. Langer G.G., Cohen S.X., Perrakis A., Lamzin V.S. (2008) Automated

macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* 3: 1171-1179.

94. Laskowski R.A., MacArthur M.W., Moss D.S., Thornton J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. App. Cryst* 26: 283-291.

95. Laskowski R. (2009). PDBsum new things. *Nucleic Acids Res*. 37: 355-359.

96. Lee S.G., Krishnan H.B., Jez J.M. (2014). Structural basis for regulation of rhizobial nodulation and symbiosis gene expression by the regulatory protein NolR. *Proc. Nat. Acad. Sci. USA* 111: 6509-6514.

97. Leslie A.G.W., Powell H.R. (2007). Processing Diffraction Data with Mosflm. Read R.J. i Sussman J.L. (ed.) Evolving Methods for Macromolecular Crystallography 245: 41-51.

98. Levy J., Bres C., Geurts R., Chalhoub B., Kulikova O., Duc G., Journet E.P., Ane J.M., Lauber E., Bisseling T., Denarie J., Rosenberg C., Debelle F. (2004). A putative Ca^{2+} and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* 303: 1361-1364.

99. Limpens E., Franken C., Smit P., Willemse J., Bisseling T., Geurts R. (2003). LysM domain receptor kinases regulating rhizobial nod factor-induced infection. *Science* 302: 630-633.

100. Lubkowski J., Dauter M., Aghaiypour K., Wlodawer A., Dauter Z. (2003). Atomic resolution structure of *Erwinia chrysanthemi* L-asparaginase. *Acta Cryst*. D59: 84-92.

101. Madsen E.B., Madsen L.H., Radutoiu S., Olbryt M., Rakwalska M., Szczyglowski K., Sato S., Kaneko T., Tabata S., Sandal N., Stougaard J. (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* 425: 637-640.

102. Malvern Instruments Ltd (2013). Zetasizer Nano user manual.

103. Maroti G., Kondorosi E. (2014). Nitrogen-fixing *Rhizobium*-legume symbiosis: are polyploidy and host peptide-governed symbiont differentiation general principles of endosymbiosis? *Frontiers in Microbiology* 5: 326.

104. McCoy A.J., Grosse-Kunstleve R.W., Adams P.D., Winn M.D., Storoni L.C., Read R.J. (2007). Phaser crystallographic software. *J. Appl. Cryst*. 40: 658-674.

105. McWilliam H., Li W., Uludag M., Squizzato S., Park Y.M., Buso N., Cowley A.P., Lopez R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acid Res.* 41: 597-600.

106. Michalska K., Brzezinski K., Jaskólski M. (2005). Crystal structure of isoaspartyl aminopeptidase in complex with L-aspartate. *J. Biol. Chem*. 280: 28484-28491.

107. Michalska K., Bujacz G., Jaskólski M. (2006). Crystal structure of plant

asparaginase. *J. Mol. Biol*. 360: 105-116.

108. Michalska K., Jaskólski M. (2006). Structural aspects of L-asparaginases, their friends and relations. *Acta Biochim. Pol*. 53: 627-640.

109. Michalska K. (2007). Badania krystalograficzne i biochemiczne amidohydrolaz. *Praca Doktorska, Wydział Chemii Uniwersytetu im. A. Mickiewicza w Poznaniu.*

110. Michalska K., Borek D., Hernandez-Santoyo A., Jaskólski M. (2008-A). Crystal packing of plant-type L-asparaginase from *Escherichia coli*. *Acta Cryst.* D64: 309-320.

111. Michalska K., Hernandez-Santoyo A., Jaskólski M. (2008-B). The mechanism of autocatalytic activation of plant-type L-asparaginases. *J. Biol. Chem*. 283: 13388-13397.

112. Miroux B., Walker J.E. (1996). Over-production of proteins in *Escherichia coli*: mutant host that allow synthesis of some membrane proteins and globular proteins at high levels. *J.Mol.Biol*. 260: 289-298.

113. Moore J.T., Uppal A., Maley F., Maley G.F. (1993). Overcoming inclusion body formation in a high-level expression system. *Protein Expr. Purif.* 4: 160-163.

114. Moreno-Enriquez A., Evangelista-Martinez Z., Gonzalez-Mondragon E.G., Calderon-Flores A., Arrerguin R., Perez-Rudea E., Huerta-Saquero A. (2012). Biochemical characterization of recombinant L-asparaginase (AnsA) from *Rhizobium etli*, a member of an increasing rhizobial-type family of L-asparaginases. *J. Microbiol. Biotechnol.* 22: 292-300.

115. Moscatiello R., Squartini A., Mariani P., Navazio L. (2010). Flavonoid-induced calcium signalling in *Rhizobium leguminosarum* bv. *viciae*. *New Phytol*. 188: 814-823.

116. Muller P., Köpke S., Scheldrick G.M. (2003). Is the bond valence method able to identify metal atoms in protein structures? *Acta Cryst*. D59: 32-37.

117. Mullis K., Faloona F., Scharf S., Saiki R., Horn G., Erlich H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 51: 263-273

118. Murray J.D. (2011). Invasion by invitation: Rhizobial infection in legumes. *Mol. Plant Microbe Interact.* 24: 631-639.

119. Murshudov G.N., Skubák P., Lebedev A.A., Pannu N.S., Steiner R.A., Nicholls R. A., Winn M.D., Long F., Vagin A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst*. D67: 355–367.

120. Nomme J., Su Y., Konrad M., Lavie A. (2012) Structures of apo and product-bound human L-asparaginase: insights into the mechanism of autoproteolysis and substrate hydrolysis. *Biochem.* 51: 6816-6826.

121. Nomme J., Su Y., Lavie A. (2014). Elucidation of the specific function of the

conserved threonine triad responsible for human L-asparaginase autocleavage and substrate hydrolysis. *J. Mol. Biol.* 426: 2471-2485.

122. Oinonen C., Tikkanen R., Rouvinen J., Peltonen L. (1995). Three-dimensional structure of human lysosomal aspartylglucosaminidase. *Nature Struct. Biol*. 2: 1102-1108.

123. Oldroyd G.E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev Microbiol*. 11:252-263.

124. Ortuno-Olea L., Duran-Vargas S. (2000). The L-asparagine operon of *Rhizobium etli* contains a gene encoding an atypical asparaginase. *FEMS Microbiol. Lett*. 189: 177-182.

125. Otwinowski Z., Minor W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* 276: 307-326.

126. Painter J., Merritt E.A. (2006). TLSMD web server for the generation of multi-group TLS models. *J. Appl. Cryst*. 39: 109-111.

127. Perret X., Staehelin C., Broughton W.J. (2000). Molecular basis of symbiotic promiscuity. *Microb. Molecular. Biol. Rev*. 64: 180-201.

128. Popp C., Ott T. (2011). Regulation of signal transduction and bacterial infection during root nodule symbiosis. *Curr. Opin. Plant Biol*. 14:458-467.

129. Pourhossein M., Korbekandi H. (2014). Cloning, expression, purification and characterisation of *Erwinia carotovora* L-asparaginase in *Escherichia coli*. *Adv. Biomed. Res.* 3:82.

130. Prahl A., Pazgier M., Hejazi M., Lockau W., Lubkowski, J. (2004). Structure of the isoaspartyl peptidase with L-asparaginase activity from *Escherichia coli*. *Acta Cryst.* D60: 1173-1176.

131. Prive G.G. (2007). Detergents for the stabilization and crystallization of membrane proteins. *Methods* 41: 388-397.

132. Pueppke S.G., Broughton W.J. (1999). *Rhizobium* sp. NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host-ranges. *Mol. Plant Microbe Interact.* 12: 293-318.

133. Quesada-Vincens D., Fellay R., Nassim T., Viprey V., Burger U., Prome J.C., Broughton W.J., Jabbouri S. (1997). *Rhizobium* sp. NGR234 NodZ protein is a fucosyltransferase. *J. Bacteriol*. 179: 5087-5093.

134. Quesada-Vincens D., Hanin M., Fellay R., Broughton W.J., Jabbouri S. (1998). *In vitro* sulfotransferase activity of NoeE, a nodulation protein of *Rhizobium* sp. NGR234. *Mol. Plant Microbe Interact*. 11: 592-600.

135. Radutoiu S., Madsen L.H., Madsen E.B., Felle H.H., Umehara Y., Grønlund M., Sato S., Nakamura Y., Tabata S., Sandal N., Stougaard J. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425: 585-

592.

136. Ritsema T., Wijfjes A.H.M., Lugtenberg B.J.J., Spaink H.P. (1996). *Rhizobium* nodulation protein NodA is a host-specific determinant of the transfer of fatty acids in nod factor biosynthesis. *Mol. Gen. Genet*. 25: 44-51.

137. Ritsema T., Lugtenberg B.J.J., Spaink H.P. (1997). Acyl-acyl carrier protein is a donor of fatty acids in the NodA-dependent step in biosynthesis of lipochitin oligosaccharides by rhizobia. *J. Bacteriol*. 179: 4053-4055.

138. Ritsema T., Gehring A.M., Stuitie A.R., van der Drift K.M.G.M., Dandal I., Lambalot R.H., Walsh C.T., Thomas-Oates J.E., Lugtenberg B.J.J., Spaink H.P. (1998). Functional analysis of an interspecies chimera of acyl carrier proteins indicates a specialized domain for protein recognition. *Mol. Gen. Genet*. 257: 641-648.

139. Rohrig H., Schmidt J., Wieneke U., Kondorosi E., Barlier I., Schell J., John M. (1994). Biosynthesis of lipooligoscacharide nodulation factors: *Rhizobium* NodA protein is involved in *N*-acylation of the chitooligosaccharide backbone. *Proc. Nat. Acad. Sci. USA* 91: 3122-3126.

140. Sambrook J., Russel D.W. (2001). Molecular cloning, a laboratory manual. *Cold Spring Harbour Laboratory Press, New York, USA.*

141. Santi C., Bogusz D., Franche C. (2013). Biological nitrogen fixation in non-legume plants. *Annals of Botany* 111:743-767.

142. Satow Y., Cohen G.H., Padlan E.A., Davies D.R. (1986). Phosphocholine binding immunoglobulin Fab McPC603: An X-ray diffraction study at 2.7 A. *J. Mol. Biol.* 190: 593-604.

143. Schalk A.M., Lavie A. (2014). Structural and kinetic characterization of guinea pig L-asparaginase type III. *Biochem.* 53: 2318-2328.

144. Schalk A.M., Nguyen H.A., Rigouin C., Lavie A. (2014). Identification and structural analysis of an L-asparaginase enzyme from guinea pig with putative tumor cell killing properties. *J. Biol. Chem*. 289: 33175-33186.

145. Schein C.H., Noteborn M.H.M. (1988). Formation of soluble recombinant proteins in *Escherichia coli* is favoured by lover growth temperature. *Biotehnology* 6: 291-294.

146. Schlaman H.R.M., Okker R.J.H., Lugtenberg B.J.J. (1992). Regulation of nodulation gene expression by NodD in Rhizobia. *J. Bacteriol.* 174: 5177-5182.

147. Schlaman H.R.M., Olsthoorn M.M.A., Harteveld M., Dorner L., Djordjevic M.A., Thomas-Oates J.E., Spaink H.P. (2006). The production of species-specific highly unsaturated fatty acyl-containing LCOs from *Rhizobium leguminosarum* bv. *trifolii* is stringently regulated by *nodD* and involves the n*odRL* genes. *Mol. Plant Microbe Interact.* 19: 215-226.

148. Schmidt J., John M., Wieneke U., Krussmann H.D., Schell J. (1986). Expression

of the nodulation gene *nodA* in *Rhizobium meliloti* and localization of the gen product in the cytosol. *Proc. Nat. Acad. Sci. USA* 83: 9581-9585.

149. Schmidt J., Wingender R., John M., Wieneke U., Schell J. (1988). *Rhizobium meliloti nodA* and *nodB* genes are involved in generating compounds that stimulate mitosis of plant cells. *Proc. Nat. Acad. Sci. USA* 85: 8578-8582.

150. Segovia L., Young P.W., Martinez-Romero E. (1993). Reclassification of American *Rhizobium leguminosarum* biovar phaseoli type I strains as *Rhizobium etli sp. nov*. *Int. J. Syst. Bacteriol.* 43:374-377.

151. Sessitsch A., Ramirez-Saad H., Hardarson G., Akkermans A.D.L., de Vos W.M. (1997). Classification of Austrian rhizobia and the Mexican isolate FL27 obtained from *Phaseolus vulgaris* L. as *Rhizobium gallicum*. *Int. J. Sys. Bacteriol.* 47: 1097-1101.

152. Shearman C.A., Rossen L., Johnston A.W.B., Downie J.A. (1986). The *Rhizobium leguminosarum* nodulation gene *nodF* encodes a polypeptide similar to acyl-carrier protein and is regulated by *nodD* plus a factor in pea root exudate. *EMBO* 5: 647-652.

153. Sheldrick G.M. (2008) A short history of SHELX. *Acta Cryst.* A**64**: 112-122.

154. Shuman S. (1994). Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J. Biol. Chem*. 269: 32678-32684.

155. Smit P., Raedts J., Portyanko V., Debelle F., Gough C., Bisseling T., Geurts R. (2005). NSP1 of the GRAS protein family is essential for rhizobial nod factor-induced transcription. *Science* 308: 1789-1791.

156. Smit P., Limpens E., Geurts R., Fedorova E., Dolgikh E., Gough C., Bisseling T. (2007). *Medicago* LYK3, an entry receptor in rhizobial nodulation factor signalling. *Plant Physiol*. 145: 183-191.

157. Smyth D.S., Mrozkiewicz M.K., McGrath W.J., Listwan P., Kobe B. (2003). Crystal structures of fusion proteins with large-affinity tags. *Protein Sci*. 12: 1313-1322.

158. Sodek L., Lea P.J., Miflin B.J. (1980). Distribution and properties of a potassium-dependent asparaginase isolated from developing seeds of *Pisum sativum* and other plants. *Plant Physiol.* 65: 22-26.

159. Sorensen H.P., Mortensen K.K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli. J. Biotechnol* 115: 113-128.

160. Spaink H.P., Wijfjes A.H.M., van der Drift K.M.G.M., Haverkamp J., Thomas-Oates J.E., Lugtenberg B.J.J. (1994). Structural identification of metabolites produced by the NodB and NodC proteins of *Rhizobium leguminosarium*. *Mol. Microbiol.* 13: 821-831.

161. Spaink H.P., Kondorosi A., Hooykaas P.J.J. (1998). The *Rhizobiaceae*, Molecular biology of model plant-associated bacteria. *Kluwer Academic Publishers,*

Dordrecht, Holandia.

162. Srikhanta Y.N., Atack J.M., Beacham I.R., Jennings M.P. (2013). Distinct physiological roles for the two L-asparaginase isozymes of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 436: 362-365.

163. Stracke S., Kistner C., Yoshida S., Mulder L., Sato S., Kaneko T., Tabata T., Sandal N., Stougaard J., Szczyglowski K., Parniske M. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 417: 959-962.

164. Su Y., Karamitros C.S., Nomme J., McSorley T., Konrad M., Lavie A. (2013). Free glicyne accelerates the autoproteolytic activation of human asparaginase. *Chemistry and Biology* 20: 533-540.

165. Sudhir A.P., Dave B.R., Prajapati A.S., Panchal K., Patel D., Subramanian R.B. (2014). Characterization of a recombinant glutaminase-free L-asparaginase (ansA3) enzyme with high catalytic activity from *Bacillus licheniformis*. *App. Biochem. Biotechnol.* 174: 2504-2515.

166. Suto H., Yasuda H., Isobe Y., Sasaki M., Imai H., Tsutsui M., Oshimi K., Komatsu N., Sugimoto K. (2010). Suppression of eIF4E expression by L-asparaginase. *Acta Haematol.* 123: 215-219.

167. Swain A.L., Jaskólski M., Housset D., Rao J.K., Wlodawer A. (1993). Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used in cancer therapy. *Proc. Nat. Acad. Sci. USA* 90: 1474-1478.

168. Taylor G.L. (2010). Introduction to phasing. *Acta Cryst*. D66: 325-338.

169. Thorn A., Sheldrick G.M. (2011). ANODE: anomalous and heavy-atom density calculation. *J. Appl. Cryst.* 44: 1285-1287.

170. Tong W.H., Pieters R., Kaspers G.J.L., te Loo D.M., Bierings M.B., van den Bos C., Kollen W.J., Hop W.C., Lanvers-Kaminsky C., Relling M.V., Tissing W.J., van der Sluis I.M. (2014). A prospective study on drug monitoring of PEGasparaginase and *Erwinia* asparaginase and asparaginase antibodies in pediatric acute lymphoblastic leukemia. *Blood* 123: 2026-2033.

171. Treilhou M., Ferro M., Monteiro C., Poinsot V. (2000). Differentiation of *O*-acetyl and *O*-carbamoyl esters of *N*-acetyl-glucosamine by decomposition of their oxonium ions. Application to the structure of the nonreducing terminal residue of nod factors. *J. Am. Soc. Mass Spectrom*. 11: 301-311.

172. Udvardi M., Poole P.S. (2013). Transport and metabolism in legume-rhizobia symbioses. *Annu. Rev. Plant Biol*. 64:781-805.

173. Van Duyne G.D., Standaert R.F., Karplus P.A., Schreiber S.L., Clardy J. (1993). Atomic structures of the human immunophilin FKBP-12 complex with FK506 and rapamycin. *J. Mol. Biol.* 229: 105-124.

174. Verma S., Metha R.K., Mati P., Rohm K.H., Sonawane A. (2014). Improvement of stability and enzymatic activity by site-directed mutagenesis of *E. coli*

asparaginase II. *Biochim. Biophys. Acta.* 1844: 1219-1230.

175. Winn M.D., Ballard C.C., Cowtan K.D., Dodson E.J., Emsley P., Evans P.R., Keegan R.M., Krissinel E.B., Leslie A.G.W., McCoy A., McNicholas S.J., Murshudov G.N., Pannu N.S., Potterton E.A., Powell H.R., Read R.J., Vagin A., Wilson K.S. (2011). Overview of the CCP4 suite and current developments. *Acta. Cryst.* D67: 235-242

176. Wlodawer A., Minor W., Dauter Z., Jaskólski M. (2008). Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J*. 275: 1-21.

177. Yao M., Yasutake Y., Morita H., Tanaka I. (2005). Structure of the type I L-asparaginase from the hyperthermophilic archaeon *Pyrococcus horikoshii* at 2.16 angstroms resolution. *Acta Cryst.* D61: 294-301

178. Yun M.K., Nourse A., White S.W., Rock C.O., Heath R.J. (2007). Crystal structure and allosteric regulation of the cytoplasmic *Escherichia coli* L-asparaginase I. *J. Mol. Biol.* 369: 794-811.

179. Zheng H., Chordia M.D., Cooper D.R., Chruszcz M., Muller P., Sheldrick G.M., Minor W. (2014). Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. *Nature Protocols* 9: 156-170

180. Zhukov V., Radutoiu S., Madsen L.H., Rychagova T., Ovchinnikova E., Borisov A., Tikhonovich I., Stougaard J. (2008). The pea *sym37* receptor kinase gene controls infection-thread initiation and nodule development. *Mol. Plant Microbe Interact*. 21: 1600-1608.

181. Zuo S., Zhang T., Jiang B., Mu W. (2015). Recent research progress on microbial L-asparaginases. *Appl. Microbiol. Biotechnol.* 99: 1069-1079.