



Dr hab. Małgorzata Łobocka, prof. PAN,
Zakład Biochemii Drobnoustrojów,
Instytut Biochemii i Biofizyki PAN,
Ul. Pawińskiego 5A
02-106 Warszawa,
Tel.: 022-592-1300 lub 022-592-1303,
Fax: 022-592-2190
E-mail: loboocka@ibb.waw.pl

Warszawa, 30. 06. 2020

Recenzja rozprawy doktorskiej mgr Dawida Kościelniaka pt.:
**„Nieprogramowany poślizg transkrypcji: analiza porównawcza bakteryjnej (*Escherichia coli*)
i bakteriofagowej (T7) polimerazy RNA"**

Promotor: dr hab. Marian Sętkas, prof. Uniwersytetu Gdańskiego (Katedra Mikrobiologii, Wydziału
Biologii Uniwersytetu Gdańskiego)

Postęp w technologiach sekwencjonowania DNA, RNA i białek znacznie poszerzył możliwości badań nad wewnątrzkomórkowym zróżnicowaniem produktów ekspresji poszczególnych genów u mikroorganizmów oraz nad jego przyczynami. Jedną z tych przyczyn są błędy w transkrypcji homopolimerowych sekwencji obecnych w matrycowym DNA, szczególnie tzw. ciągów kilku A lub T, stosunkowo częstych w genomach dużej liczby bakterii. Choć mogą one powodować fenotypowe odwrócenie efektów delekcji lub insercji w genach kodujących białka i są problemem w ekspresji genów w bakteriach na potrzeby przemysłu, mechanizmy ich powstawania oraz uwarunkowania w obrębie sekwencji DNA, na matrycy których powstają są słabo poznane.

Celem przedstawionej mi do oceny pracy doktorskiej mgr Dawida Kościelniaka, było porównanie nieprogramowanych poślizgów transkrypcji dwóch polimeraz RNA o odmiennej strukturze: wielopodjednostkowej polimerazy RNA *Escherichia coli* reprezentującej polimerazy bakteryjne oraz jednopodjednostkowej polimerazy DNA bakteriofaga T7 reprezentującej polimerazy fagowe, ze szczególnym uwzględnieniem wpływu długości i typu samych sekwencji homopolimerowych oraz sekwencji otaczających, a także niektórych czynników białkowych wpływających na elongację, na częstość poślizgów.

Praca została wykonana w Katedrze Mikrobiologii, Wydziału Biologii Uniwersytetu

Gdańskiego, pod kierunkiem prof. Mariana Sęktasa, w którego kręgu zainteresowań od kilku lat są poślizgi polimeraz RNA na homopolimerowych sekwencjach oraz uwarunkowania wpływające na ten proces. Wybór do równoległych badań w podobnych warunkach dwóch polimeraz o odmiennej strukturze, różniących się ponadto zdolnościami korektorskimi istotnie podwyższa wagę otrzymanych wyników, tym bardziej, że autor zadbał tam gdzie było to możliwe o ujednoczenie zarówno efektywności transkrypcji przez obie polimerazy w badaniach porównawczych jak i systemów reporterowych pozwalających na jakościowe i ilościowe porównanie otrzymanych wyników, co wymagało wielu żmudnych badań wstępnych. Warto podkreślić, że w prezentowanej pracy po raz pierwszy porównano zdolności polimerazy RNA *E. coli* i bakteriofaga T7 do poślizgu podczas transkrypcji homopolimerów A lub T krótszych niż 9 reszt nukleotydowych. W połączeniu z wysokim poziomem standaryzacji warunków działania obu polimeraz pozwoliło to na zaobserwowanie istotnych różnic w ich zdolności do poślizgu podczas transkrypcji, a tym samym częstości generowania polimorficznych transkryptów genów zawierających homopolimery A lub T i różnej zależności tej częstości od typu, długości i bezpośredniego otoczenia sekwencji homopolimerowych, a także różnic w preferowanych zmianach typu insercji lub delecji powstających z udziałem każdej z badanych polimeraz.

Kluczowe dla możliwości badania wpływu sekwencji poli-A i poli-T na częstości poślizgów poszczególnych polimeraz podczas transkrypcji było posłużenie się odpowiednimi systemami wskaźnikowymi. Autor przygotował i wykorzystał kilka. Do pomiaru częstości poślizgu -1 i +1 posłużył się genem *mboIIM2* kodującym metylotransferazę *Moraxella bovis*, zawierającym kilka sekwencji homopolimerowych i sklonowanym w plazmidzie zarówno w wariacie pod kontrolą promotora dla polimerazy RNA *E. coli* jak i dla polimerazy RNA faga T7. Dobrał stężenia induktorów każdego z promotorów tak, by choć w części skompensować różnice w elongacji transkrypcji pomiędzy obydwoma polimerazami. Wariant genu *mboIIM2* z delecją jednej reszty adenozyliny w sekwencji poli-A wykorzystał do pomiaru poślizgu polimerazy o -1 poprzez pomiar zarówno aktywności metylującej białka o fenotypie dzikim tworzonego na skutek poślizgu jak i względnego poziomu przypuszczalnego białka o fenotypie dzikim i białka mutantu o masie mniejszej od białka dzikiego. Dużą część pracy zajmuje opis badań zmierzających do oddzielenia efektów związanych z poślizgiem polimerazy od towarzyszącego im efektu polarności transkrypcji, powodującego znacznie zmniejszenie ilości transkryptów dystalnej części genu reporterowego u mutantu z przedwczesnym sygnałem terminacji translacji w tym genie. Do badań porównawczych wpływu różnych czynników na efekt polarności transkrypcji autor wykorzystał pomiary ilości mRNA reprezentujących poszczególne części transkryptu z wykorzystaniem odwrotnej transkrypcji i ilościowej reakcji PCR w czasie rzeczywistym z parami starterów komplementarnych do różnych

części badanego genu. Pozwoliło to na wykazanie, że efekt polarności transkrypcji obserwowany u mutantu z delecją reszty adenozyiny w rejonie poli-A genu *mboIIM2* powiązany jest z degradacją pozbawionego rybosomów mRNA przez RNazę E i nie zależy od czynnika terminacyjnego Rho. Brak udziału czynnika Rho w tym procesie autor wykazał również przez zademonstrowanie, że bicyclomycyna, blokująca działanie Rho nie ma wpływu na obserwowany efekt polarności transkrypcji.

Badania z wykorzystaniem genów reporterowych umożliwiły ocenę częstości poślizgów o -1 na sekwencji poli-A, ale nie dawały pełnego obrazu poślizgów polimerazy. Dlatego autor porównywał pojedyncze odczyty sekwencji produktów odwrotnej transkrypcji otrzymanych z wykorzystaniem puli transkryptów genu *mboIIM2* z delecją reszty adenozyiny w ciągu poli-A, wytworzonych przez polimerazę RNA *E. coli* i T7. Pozwoliło to na ocenę częstości delecji i insercji w rejonach homopolimerowych mRNA wytworzonego przez każdą z tych polimeraz i pokazanie różnic pomiędzy polimerazami w częstości poślizgów, preferencji dla poszczególnych rejonów poślizgowych i preferowanych zmian wynikających z poślizgów. Mam dodatkowe pytanie do tej części wyników. Polimerazy DNA używane do sekwencjonowania też w warunkach *in vitro* myślą się na sekwencjach poślizgowych. Jaka więc może być w eksperymencie sekwencjonowania cDNA otrzymanego na matrycach polimorficznych mRNA kontrola określająca ile z obserwowanych w otrzymanych sekwencjach DNA polimorfizmów reprezentuje te przepisane z mRNA, a ile te wygenerowane przez polimerazy użyte do sekwencjonowania?

Ciekawe obserwacje przedstawione przez autora dotyczą wpływu antyterminatorowego białka N i wiążącej to białko sekwencji *nutL* na częstość poślizgów polimerazy RNA *E. coli* na krótkich sekwencjach poli-A. Mimo, że w poprzednich badaniach wykazano wpływ N i *nutL* na częstość poślizgów polimerazy na sekwencjach poli-A i poli-T dłuższych niż 8 reszt nukleotydowych, i autor potwierdził te obserwacje, wykazał jednocześnie, że w przypadku krótszych sekwencji poli-A i poli-T, brak jest wpływu N/*nutL* na częstość poślizgów, a białko N działa w tym wypadku nie jako antyterminator, a raczej jako pozytywny czynnik elongacyjny. Zwiększa więc stężenie dystalnej części transkryptów takich sekwencji. Ciekawa jestem, czy autor mógłby zaproponować z czego może wynikać różnica wpływu białka N i sekwencji *nutL* na częstość poślizgów polimerazy RNA w rejonach długich i krótkich sekwencji homopolimerowych poli-A lub poli-T.

Inny system reporterowy został wykorzystany przez autora w badaniach nad preferencjami polimerazy RNA *E. coli* i T7 do poślizgu w zależności od rodzaju sekwencji homopolimerowej, jej długości i otoczenia nukleotydowego. W tym celu autor wykorzystał geny fuzyjne składające się z genu reporterowego *gfp* poprzedzonego fragmentami DNA zawierającymi sekwencje poli-A lub

poli-T różnej długości i w różnym bezpośrednim otoczeniu nukleotydowym. Przez dobór sekwencji o odpowiedniej długości poprzedzających *gfp* i poza ramką odczytu o -1 lub +1 możliwe było badanie efektów poślizgu przywracających syntezę fuzji z *gfp* wykazującej fluorescencję. Jednak uzyskanie wiarygodnych wyników wiązało się z koniecznością pokonania szeregu trudności technicznych związanych z koniecznością wyboru spośród badanych fragmentów genu *mboIIM2* z delecjami lub insercjami takich, których naprawione produkty nie zagłuszały fluorescencji w fuzji z białkiem Gfp, potwierdzenia, że sekwencja otrzymanych fuzji odpowiada białkom powstałym po poślizgu o -1 lub +1, dodatkowych badań nad wpływem zmian nukleotydowych powstałych w wyniku poślizgu na początku genu fuzyjnego na poziom ekspresji tego genu i tym samym poziom fluorescencji, a także dodatkowej normalizacji wyników pomiaru fluorescencji zależnie od liczby reszt aminokwasowych z N-końca białka Gfp pozostawionych na skutek poślizgu. Dopiero po tak licznych badaniach wstępnych autor mógł dokonać właściwych analiz pozwalających m. in. na określenie minimalnej długości sekwencji poli-A i poli-T niezbędnych do indukcji poślizgu przez każdą z badanych polimeraz oraz preferencji dla określonych reszt nukleotydowych w miejscach obserwowanych poślizgów i w ich bezpośrednim sąsiedztwie. W analizie wyników wykorzystał globalne pomiary fluorescencji hodowli znormalizowanych pod względem liczby komórek, przyżyciowe obserwacje fluorescencji w pojedynczych komórkach, oraz pomiary ilości różnych form białek fuzyjnych w komórkach metodą Western blotting z wykorzystaniem przeciwciał anti-Gfp. W dalszej części badań zbadał wpływ białek GreA i GreB znanych jako czynniki antypauzowe na efektywność poślizgu polimerazy RNA *E. coli*. Wyniki tych eksperymentów wskazały na negatywny wpływ tych białek na częstość poślizgu, zamiast spodziewanego wpływu pozytywnego. Ciekawa jestem, jak możnaby ten wynik zinterpretować.

Ogólnie wyniki badań zaprezentowanych przez autora pokazały olbrzymie bogactwo czynników wpływających na poślizgi polimeraz RNA podczas transkrypcji rejonów poli-A i poli-T, oraz duże różnice w tym względzie pomiędzy wielopodjednostkową polimerazą RNA *E. coli* i jednopodjednostkową polimerazą bakteriofaga T7. Polimeraza RNA faga T7 wykazała przy tym dużo większe tendencje do poślizgu, nawet podczas transkrypcji krótkich sekwencji homopolimerowych. Dodatkowo w pracy wykazano, że różnice pomiędzy obydwoma polimerazami w efektywności tworzenia mRNA nieprawidłowo przetłumaczonych na skutek poślizgu są potęgowane przez tzw. efekt polarności transkrypcji związany z przedwczesną degradacją gołych transkryptów tworzonych przez polimerazę RNA *E. coli*, przez RNazę E. Nie jest dla mnie jasne, dlaczego efekt polarny transkrypcji, jak podkreślił Autor, nie występuje w przypadku polimerazy RNA faga T7. Czy wynika to z braku powinowactwa czynników odpowiedzialnych za ten efekt do polimerazy RNA T7?

Ciekawe wnioski wynikają z przeprowadzonej przez Autora analizy genomów *E. coli* i bakteriofaga T7 pod kątem określenia zawartości naturalnych sekwencji poli-A i poli-T w genach, oraz rejonach międzygenowych i promotorowych w tych genomach. Być może to duża skłonność polimerazy RNA faga T7 do poślizgów ogranicza liczbę długich ciągów homopolimerowych sekwencji w genomie tego faga w porównaniu z genomem *E. coli*. Z drugiej strony analiza częstości występowania ciągów homopolimerowych w genomie *E. coli* wykazała istotną nadreprezentację ciągów poli-A/poli-T o długości 5-9 reszt nukleotydowych, co wydaje się być zakonserwowaną cechą genomów wielu bakterii. Przy omawianiu tej obserwacji warto uwzględnić znaczenie właśnie takich ciągów dla pakowania DNA w nukleoidzie bakteryjnym (Tolstorukov i in. 2005, NAR, 33, 3907-18).

Rozprawa, łącznie z liczącym ponad 300 pozycji spisem literatury liczy 202 strony tekstu i jest przygotowana bardzo starannie. Dodatkowo tekst jest wzbogacony o materiały uzupełniające. W zawartym na 39 stronach wstępie autor skoncentrował się na opisie struktury dwóch porównywanych w pracy polimeraz RNA oraz ogólnym opisie poszczególnych etapów transkrypcji. Moim zdaniem zabrakło wprowadzenia w zagadnienia dotyczące czynników białkowych wpływających na efektywność i wierność elongacji transkrypcji i o mechanizmach ich działania, tym bardziej, że w doświadczeniach opisanych w pracy badany był wpływ kilku z tych czynników (np. GreA i GreB), na nieprogramowane poślizgi polimerazy RNA *E. coli*. Nie poruszono też problemu zahamowania transkrypcji jako części mechanizmu korekty błędów transkrypcyjnych. Chociaż wyniki ostatnich badań potwierdzają tylko ograniczoną rolę GreA w tym procesie, dodanie we wstępie informacji o GreA i GreB oraz innych białka uwzględnionych w prezentowanych badaniach pomogłoby w lepszym zrozumieniu dalszych części pracy. Najciekawsza część wstępu poświęcona jest przykładom występowania poślizgów transkrypcji w naturze oraz ich konsekwencjom, co podkreśla ważność tematyki podjętej przez autora pracy. Zawarty na 45 stronach tekstu opis materiałów i stosowanych metod jest na tyle precyzyjny, że powtórzenie badań opisanymi metodami nie powinno sprawiać problemu. Na uwagę zasługuje zróżnicowany dobór wskaźników do badania poślizgów polimeraz RNA podczas elongacji transkrypcji, włączając zarówno pomiar ilości RNA jak i pomiar ilości różnych wariantów białek z wykorzystaniem metody Western blotting oraz fluorescencji fuzji z białkiem Gfp. Dla pokazania wpływu uwarunkowań sekwencji DNA na częstość poślizgów i rodzaj generowanych przez nie mutacji autor skonstruowała aż 116 nowych plazmidów oraz szereg mutantów *E. coli*. W opisie wyników zwraca uwagę staranność w doborze systemów detekcji poślizgów polimerazy, tak by wyeliminować w jak największym stopniu różnice albo z nimi niezwiązane, albo będące tylko ich skutkiem pośrednim. W mojej opinii praca jest unikalnym tak dokładnym i przemyślanym studium

porównawczym polimerazy RNA *E. coli* i polimerazy RNA bakteriofaga T7, pod kątem zdolności obu polimeraz do poślizgów oraz ich preferencji w tym zakresie. Przedstawione dane nie wyczerpują tematu związanego z poślizgami obu polimeraz podczas transkrypcji sekwencji homopolimerowych, jednak wskazują na istotne różnice pomiędzy polimerazami, pokazują szereg czynników istotnych dla częstości poślizgów i dają porządne podstawy dla kontynuacji badań nad poślizgami z utrzymaniem wskazanych w pracy wysokich standardów kontroli. Pokazują też jednoznacznie, że heterogenność transkryptów związana z poślizgami polimeraz przy transkrypcji genów zawierających sekwencje homopolimerowe powinna być brana pod uwagę przy wykorzystywaniu każdej z tych polimeraz do ekspresji sklonowanych genów w procesach produkcji białek. Wiele z tych istotnych zagadnień zostało poruszonych przez autora w dyskusji wyników pracy. Wynika z niej jasno, że dopiero badania poślizgu polimeraz z uwzględnieniem wpływu na transkrypcję szeregu innych czynników, takich jak np. sprzężenie transkrypcyjno-translacyjne, terminacja transkrypcji, degradacja RNA oraz regulacja pauzowania, czy też i szybkości oraz procesywności samych polimeraz mogą dać wiarygodny obraz tego procesu i częstości jego zachodzenia.

Z obowiązku recenzenta mam kilka drobnych uwag i pytań odnośnie tekstu pracy.

1. Na str. 16 wstępu znalazłam dotyczące polimerazy RNA faga T7 sformułowanie "Nie transkrybuje ona genów obsługiwanych przez naturalne promotory". Moim zdaniem właściwsze byłoby napisanie "będących pod kontrolą naturalnych promotorów".
2. Str. 25, podpis pod Ryc. 7: Co autor miał na myśli używając sformułowania "drugorzędowy kanał"?
3. W opisie Materiałów, w Tabeli 2, faga P1vir opisano jako mutanta z delecją rejonu *attP*. Tymczasem fag P1 nie ma miejsca *attP*, bo jego profag występuje w komórce w postaci plazmidu, a nie jako DNA zintegrowane z chromosomem. Przyczyna niezdolności mutanta P1vir do lizogenii jest inna niż opisano w pracy.
4. Str. 62: podrozdział 4.4.2.19" 15 mM MgSO₄ nie jest buforem.
5. Str. 72: Zamiast sformułowania "Transformacja plazmidowa" właściwiej byłoby napisać "Transformacja plazmidowym DNA", a zamiast tytułu rozdziału "Indukcja hodowli bakteryjnej" dla jasności i poprawności można było napisać o indukcję czego w komórkach chodzi.
6. Na str. 89 i 93 autor użył sformułowania "cofnięte końce". Czy nie lepiej byłoby zamiast tego użyć powszechnej w literaturze nazwy "lepkie końce"?
7. Opis konstrukcji plazmidu pBADT7*mboIIM2ΔA356* pojawia się nieoczekiwanie na końcu rozdziału Metody. Czy nie lepiej byłoby zamieścić ten opis razem z opisami konstrukcji

innych plazmidów?

8. Str. 96: Na jakiej podstawie dobrano akurat takie stężenia induktorów dla promotora RNAP T7 i dla promotora RNAP *E. coli*. Czy robiono pomiary poziomu transkryptów przy różnych stężeniach induktorów, czy też posłużono się danymi literaturowymi?
9. Str. 107: Zamiast "81 nukleotydów" i "29 aminokwasów" poprawnie powinno być "81 reszt nukleotydowych" i "29 reszt aminokwasowych"
10. Str. 112: Czy chodziło o "Wpływ szybkości i procesywności", a nie "Wpływ szybkości procesywności"?
11. Str. 113: Nie zgadzam się z interpretacją wyników przedstawionych na Ryc. 26. Wg tej interpretacji w przypadku włączonego do chromosomu genu *mboITM2ΔA356* zarówno z udziałem polimerazy RNA *E. coli* jak i z udziałem polimerazy T7 następuje produkcja zarówno krótszej jak i długiej formy białka i jest efektywniejsza w przypadku polimerazy RNA T7. Dane przedstawione na rysunku pokazują jednak, że długa forma dominuje po indukcji transkrypcji z udziałem polimerazy RNA T7, a krótka po indukcji transkrypcji z udziałem polimerazy RNA *E. coli*.

Przedstawione uwagi w niczym nie umniejszają wartości pracy, którą oceniam wysoko. Autor osiągnął zamierzony cel. Ponadto, główne wyniki pracy zostały opublikowane w dwóch publikacjach w międzynarodowym czasopiśmie *Microbial Cell Factories* o współczynniku oddziaływania (IF) wynoszącym (4,402) i były w związku z tym już poddane wnikliwej ocenie przez specjalistów. Mgr Kościelniak jest pierwszym z czterech współautorów w jednej z tych prac i drugim w drugiej. Podsumowując stwierdzam, że recenzowana praca spełnia wymogi formalne i merytoryczne stawiane rozprawom doktorskim i odpowiada warunkom określonym ustawie o tytule i stopniach naukowych. Przedstawione wyniki stanowią istotny wkład do wiedzy na temat nieprogramowanych poślizgów dwóch różnych polimeraz RNA na sekwencjach homopolimerowych podczas transkrypcji. Stawiam wniosek do Rady Dyscypliny Nauk biologicznych Uniwersytetu Gdańskiego, o przyjęcie rozprawy i dopuszczenie mgr Dawida Kościelniaka do dalszych etapów przewodu doktorskiego. Jednocześnie, ze względu na wysoką wartość merytoryczną pracy, wnoszę o jej wyróżnienie stosowną nagrodą.

