



**Uniwersytet
Gdański**

Wydział Biologii
Uniwersytetu Gdańskiego

mgr Aleksandra Wiśniewska

Molekularny mechanizm krzyżowego
działania dwóch czynników
transkrypcyjnych: białka C systemu
restrykcyjno-modyfikacyjnego Csp231I oraz
represora RacR defektywnego profaga
Escherichia coli.

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

Promotor: dr hab. Iwona Mruk, prof. UG
Katedra Mikrobiologii

GDAŃSK 2023

Tytuł pracy w języku angielskim:

Molecular basis for the crosstalk between two unrelated bacterial transcription factors – the regulatory C protein of a restriction-modification system Csp231I and RacR repressor of a defective prophage in *Escherichia coli* cells.



NARODOWE CENTRUM NAUKI

Źródła finansowania

Grant OPUS10 nr 2015/19/B/NZ2/01835 finansowany przez Narodowego Centrum Nauki pt. „Dynamika kontroli ekspresji genów kodujących bakteryjne systemy restrykcyjno-modyfikacyjne w czasie z zastosowaniem technologii monitorowania pojedynczych komórek”.

Grant OPUS18 nr 2019/35/B/NZ2/00701 finansowany przez Narodowego Centrum Nauki pt. „Molekularne podstawy zakłócania transkrypcji poprzez tzw. cross-talk czynników transkrypcyjnych, i jego wpływ na żywotność bakterii”.

Składam najserdeczniejsze podziękowania Pani dr hab. Iwonie Mruk, prof. UG za nieocenioną pomoc na wszystkich etapach powstawania tej pracy, zaangażowanie, przekazaną wiedzę i poświęcony czas.

Chciałabym również podziękować mojemu mężowi. Arturze, za ogromne wsparcie, wyrozumiałość i cierpliwość.

Dziękuję Mamie i Tacie. Kochani Rodzice, to dla Was i dzięki Wam.

Spis treści

Spis treści	5
Wykaz stosowanych skrótów	9
Streszczenie.....	10
Abstract.....	12
1. Wstęp.....	14
1.1. Transkrypcja	14
1.2. Czynniki sigma (σ)	14
1.3. Czynniki transkrypcyjne	15
1.4. Globalne białka regulatorowe w transkrypcji.....	16
1.5. Miejscowi regulatorzy transkrypcji	17
1.6. Bakteryjne sieci regulacji transkrypcji.....	19
1.7. Zakłócenie bakteryjnych sieci regulacji transkrypcji	21
1.8. Zakłócenie sieci globalnych regulatorów transkrypcji.....	23
1.9. Zakłócenie sieci miejscowych regulatorów transkrypcji.....	25
1.10. Różne „oblicza” zjawiska reakcji krzyżowej (ang. <i>cross-talk</i>).....	26
1.11. Funkcje systemów R-M	26
1.12. Systemy R-M Typu II	27
1.13. Organizacja genetyczna systemu R-M Csp231I.....	28
2. Cel pracy	30
3. Materiały.....	31
3.1. Szczepy bakteryjne	31
3.2. Plazmidy bakteryjne	31
3.3. Oligonukleotydy i substraty DNA	33
3.4. Wzorce masowe	36
3.5. Antybiotyki	36
3.6. Podłoża wzrostowe	36
3.6.1. Pożywka płynna LB (Luria Broth), pH 7,2	36
3.6.2. Pożywka stała LA (Luria Agar), pH 7,2	36
3.6.3. Pożywka minimalna M9 (skład na 100 ml)	37
3.7. Bufory i żele do elektroforez	37
3.7.1. Elektroforeza DNA w żelu agarozowym	37
3.7.1.1. Bufor elektrodowy TBE	37
3.7.1.2. Żel agarozowy 0,8%	37
3.7.1.3. Barwnik obciążający do elektroforezy agarozowej 6 × stężony	37

3.7.2. Elektroforeza białek w żelu poliakrylamidowym	37
3.7.2.1. Bufor SDS-PAGE pH 8,3	37
3.7.2.2. Bufor Tris-HCl pH 8,8 (dolny).....	38
3.7.2.3. Bufor Tris-HCl pH 6,8 (górnny)	38
3.7.2.4. Roztwór 30% akrylamidów	38
3.7.2.5. Żel poliakrylamidowy rozdzielający (dolny) 12,5%	38
3.7.2.6. Żel poliakrylamidowy zagęszczający (górnny) 4%	38
3.7.2.7. Barwnik Laemmli 4 × stężony	38
3.7.2.8. Barwienie żeli białkowych	38
3.7.2.9. Barwnik Coomassie Brilliant Blue (CBB)	38
3.7.2.10. Odbarwiacz.....	39
3.7.3. Elektroforeza białek w żelu tricinowy	39
3.7.3.1. Bufor elektrodowy Tris-HCl pH 8,9 (dolny).....	39
3.7.3.2. Bufor elektrodowy Tris (górnny)	39
3.7.3.3. Bufor tricinowy pH 8,45.....	39
3.7.3.4. Żel tricinowy rozdzielający (dolny) 12,5 %.....	39
3.7.3.5. Żel tricinowy zagęszczający (górnny) 4 %	39
3.8. Bufory i roztwory używane do analizy białek	40
3.8.1. Bufor 2 x NPi pH 8,0	40
3.8.2. Bufor NPi-10	40
3.8.3. Bufor NPi-20	40
3.8.4. Bufor NPi-150	40
4. Metody	41
4.1. Amplifikacja fragmentów DNA metodą PCR (reakcja łańcuchowa polimerazy ang. <i>Polymerase Chain Reaction</i>).....	41
4.2. Mutageneza miejscowo-specyficzna	41
4.3. Elektroforeza agarozowa DNA	42
4.4. Oczyszczanie fragmentu DNA z żelu agarozowego	42
4.5. Izolacja DNA z żelu metodą elektroelucji.....	42
4.6. Trawienie DNA enzymami restrykcyjnymi.....	42
4.7. Defosforylacja 5' końców DNA.....	43
4.8. Ligacja fragmentów DNA trawionych enzymami restrykcyjnymi	43
4.9. Przygotowanie komórek kompetentnych.....	43
4.10. Transformacja komórek kompetentnych	43
4.11. Nadprodukcja białka w systemie bakteriofaga T7	44
4.12. Oczyszczanie białka	44
4.12.1. Liza komórek bakteryjnych.....	44
4.12.2. Chromatografia powinowactwa na złożu TALON	44
4.12.3. Dializa	45
4.13. Pomiar stężenia białka.....	45
4.14. Sączenie molekularne	45
4.15. Elektroforeza w żelu poliakrylamidowym w warunkach denaturujących. 46	46

4.15.1. Przygotowanie żelu poliakrylamidowego	46
4.15.2. Wykonanie prób do elektroforezy białkowej	46
4.15.3. Elektroforeza białek w żelu poliakrylamidowym	46
4.16. Barwienie fluorescencyjne LIVE/DEAD	46
4.17. Przygotowanie preparatu do mikroskopii fluorescencyjnej	47
4.18. Indukcja ekspresji genów w hodowli bakteryjnej za pomocą arabinozy ...	47
4.19. Oznaczanie aktywności β-galaktozydazy (test ONPG).....	47
4.20. Oznaczanie efektywności tworzenia kolonii bakteryjnych	48
4.21. Test opóźnienia migracji DNA w żelu EMSA (ang. <i>Electrophoretic Mobility Shift Assay</i>).....	49
4.22. Metoda śladu DNA (ang. <i>DNA footprinting</i>).....	49
4.23. Relatywna restrykcja - oznaczanie efektywności tworzenia lysinek bakteriofagowych	50
4.24. Analizy bioinformatyczne.....	51
5. Wyniki.....	52
5.1. Wprowadzenie systemu R-M Csp231I do <i>E. coli</i> powoduje filamentację	52
5.2. Wpływ wariantów systemu R-M Csp231I na żywotność komórek <i>E. coli</i> MG1655.....	55
5.3. Analiza transkryptomyczna komórek niosących plazmidy z systemem R-M typu dzikiego oraz jego wariantów	56
5.4. Analiza <i>in silico</i> rejonu faga Rac	58
5.5. Badanie wpływu genu C na żywotność komórek bakteryjnych w szczepie posiadającym rejon profaga Rac w genomie.....	60
5.6. Mikroskopia fluorescencyjna LIVE/DEAD komórek bakteryjnych MG1655WT niosących białko C na plazmidzie	63
5.7. Badanie fenotypu filamentacyjnego w czasie dla komórek z aktywnym genem regulatora C.....	64
5.8. Badanie ekspresji białka FtsW w komórkach filamentujących	66
5.9. Wpływ białka C na ekspresję genów <i>racR</i>, <i>ydaS</i> oraz <i>ydaT</i>.....	69
5.10. Czynniki transkrypcyjne i ich wpływ na ekspresję genów operonu RacR.	73
5.11. Badanie zdolności wiązania się białka C do rejonu <i>rac</i>.....	74
5.12. Ustalenie miejsc wiązania białka C do sekwencji nukleotydowej rejonu faga kryptycznego Rac z użyciem metody śladu DNA (ang. <i>DNA footprinting</i>) .	76
5.13. Odziaływanie białka RacR z rejonem <i>rac</i>.....	81
5.13.1. Nadprodukcja i oczyszczanie białka RacR	81
5.13.2. Badanie interakcji białka RacR z rejonem faga <i>rac</i>	84
5.14. Odziaływanie kompetycyjne białek C i RacR w obrębie rejonu faga <i>rac</i> ..	85
5.15. Ustalenie miejsca wiązania białka RacR do sekwencji nukleotydowej rejonu faga <i>rac</i> zawierającej rejon intergenowy	86

5.16. Analiza <i>in silico</i> regionu intergenowego pomiędzy genem <i>racR</i> , a <i>ydaS</i>	88
6. Dyskusja.....	91
7. Podsumowanie.....	101
8. Literatura	102

Wykaz stosowanych skrótów

- Amp^R** – oporność na ampicylinę
APS – (ang. *ammonium persulfate*) nadsiarczan amonu
BSA – (ang. *bovine serum albumin*) surowicza albumina bydlęca
CBB – Coomassie Brilliant Blue
CFU – (ang. *colony forming unit*) jednostka tworząca kolonię
Cm^R – oporność na chloramfenikol
DNA – kwas deoksyrybonukleinowy
dNTPs – deoksynukleotydy
E. coli – *Escherichia coli*
EDTA – sól dwusodowa kwasu etylenodwuaminoczworoocowego
HTG – (ang. *horizontal gene transfer*) horyzontalny transfer genów
IPTG – izopropyl-β-D-tiogalaktopiranozyd
IGR – rejon intergenowy
Kan^R – oporność na kanamycynę
kDa – kilodalton
KM – Katedra Mikrobiologii
lacZ – gen kodujący β-galaktozydazę
LacZ – produkt genu *lacZ*, białko β-galaktozydaza
MTaza – metylotransferaza DNA
NAP – (ang. *nucleoid-associated proteins*) białka związane z nukleoidem
obr. /min. – (ang. *revolutions per minute*, rpm) obroty na minutę
OD – (ang. *optical density*) gęstość optyczna
PAGE – (ang. *polyacrylamide gel electrophoresis*) elektroforeza w żelu poliakrylamidowym
p.z. – par zasad
PCR – reakcja łańcuchowa polimerazy (ang. *polimerase chain reaction*)
PFU – jednostka tworząca łysinkę (ang. *plaque-forming unit*)
REaza – endonukleaza restrykcyjna
R-M – system restrykcyjno-modyfikacyjny
Ryc. – rycina
SD – (ang. *standard deviation*) odchylenie standardowe
SDS – sól sodowa siarczanu dodecylu
Tet^R – oporność na tetracyklinę
TEMED – N, N, N', N'- tetrametyloetylenodiamina
Tris – trójhydroksymetyloaminometan
UG – Uniwersytet Gdański
WT – szczep typu dzikiego

Streszczenie

Podstawowe procesy życiowe u wszystkich organizmów żywych regulowane są najczęściej poprzez kontrolę ekspresji genów, która pozwala dostosować się do szybko zmieniającego się otoczenia, co zwiększa szansę przetrwania. Taka kontrola jest wielopoziomowa, ale w dużej mierze odbywa się na etapie ekspresji genów w procesie transkrypcji. U bakterii, regulacja ekspresji genów zwykle nie odbywa się dla pojedynczych genów, ale całych grup genów (operonów/regulonów), które razem tworzą komórkową globalną sieć transkrypcyjną. Jej sprawne funkcjonowanie, wzajemne powiązanie i koordynacja procesów regulatorowych opierają się głównie na działaniu czynników transkrypcyjnych (TF, ang. *transcription factor*), które zwykle są małymi białkami wiążącymi DNA. TF oddziałują z DNA poprzez rozpoznawanie specyficznych, docelowych sekwencji nukleotydowych (ang. *target sites*; *primary sites*), ale są w stanie również rozpoznawać inne sekwencje, poza swoim celem (ang. *off-target sites*; *secondary sites*), czasem w sposób bardziej rozluźniony. Nie jest jednak jasne do tej pory, jakie mechanizmy rządzą tymi interakcjami: docelowymi i poza swoim celem, i jak one wpływają na ogólną sieć regulatorową. Ważne jest to także w sytuacji, gdy komórka pobiera fragment DNA niosący genetyczny moduł DNA zawierający TF, np. na drodze horyzontalnego transferu genów. W toku naszych badań zauważyliśmy, że wprowadzenie genów systemu restrykcyjno-modyfikacyjnego (R-M) Typu II - Csp231I na plazmidzie, skutkuje ciekawym zjawiskiem tworzenia się drastycznie wydłużonych komórek *E. coli*.

Celem badań w przedstawianej rozprawie doktorskiej było zbadanie mechanizmu molekularnego odpowiedzialnego za proces formowania się wydłużonych komórek bakteryjnych (tzw. filamentów) w szczepie *Escherichia coli* MG1655 niosących geny systemu R-M Typu II Csp231I, pochodzącego ze spokrewnionej bakterii *Citrobacter* sp. RFL231. Badania wykazały, że białko regulatorowe C, odpowiedzialne za regulację ekspresji genów systemu R-M, bierze także udział w zakłócaniu transkrypcji w regulonie *E. coli* w obrębie rejonu profaga kryptycznego Rac. Pokazano, że dochodzi do zjawiska krzyżowego działania (ang. *transcriptional cross-talk*) pomiędzy dwoma czynnikami transkrypcyjnymi: białkiem C pochodzącym z systemu R-M oraz represorem RacR defektywnego profaga Rac. Zmapowano miejsce wiązania białka C w obrębie sekwencji DNA kodującej gen *racR*. Wykazano, że to wiązanie znacząco zmniejsza ekspresję represora RacR, i ten nie jest w stanie dłużej blokować ekspresji sąsiednich genów *ydaS* oraz *ydaT*. Z kolei, ich de-represja powoduje powstawanie filamentów komórkowych

prawdopodobnie przez ujawnienie się ekspresji genu YdaT, który w warunkach fizjologicznych jest kompletnie wyciszony i nie można wykryć jego transkryptu. W pracy doktorskiej zidentyfikowano elementy genetyczne biorące udział w deregulacji operonu *racR-ydaST* przez regulator C w warunkach *in vivo* oraz *in vitro*. Zbadano wpływ tego procesu na żywotność komórek gospodarza. Nie jest jasne czy te zjawiska są przypadkowe, ale otrzymane wyniki pokazują przykład zdarzenia horyzontalnego transferu genów, które może prowadzić do zmniejszenia żywotności komórek gospodarza, a nawet jego śmierci.

Abstract

Fundamental processes in all living systems are mostly regulated at the level of gene expression, which provides adaptation to the rapidly changing environments to increase the chances of survival. This control is multi-layered, but largely operates at the level of transcription. Bacterial gene expression regulation usually impacts large groups of genes (regulons/operons), which together form global transcriptional networks. Their efficient functioning, interconnectivity and orchestration rely mainly on the action of individual DNA binding proteins called transcription factors (TFs). TFs interact not only with their specific target sites (primary sites) determined by DNA sequence motif, but also with secondary (off-target) sites, and vary in their promiscuity. It is not clear yet what mechanisms govern the interactions with secondary sites, and how such rewiring affects the overall regulatory network. This is especially crucial during transfer of a genetic module carrying a TF gene into a new host, e.g. *via* horizontal gene transfer. In the course of our research, we noticed that introduction of genes coding for Type II restriction-modification (R-M) system Csp231I led to manifestation of the abnormal cell filamentation.

The main objective of this work was to determine the molecular mechanism responsible for the formation of elongated bacterial cells (filaments) in the *Escherichia coli* MG1655 strain carrying genes of the Type II Csp231I restriction-modification system, derived from a related bacterium *Citrobacter sp.* RFL231. Our results indicated that the primary source of the cell defect was the interference of the introduced TF (C protein), as a part of a R-M system, with the host genetic network operated by the RacR repressor, an essential regulator (TF) of the cryptic Rac prophage. The C protein exerts a transcriptional cross-talk with another TF of the host, the RacR repressor. We showed that the C protein binds to its unrelated, off-target site within coding sequence of the *racR* gene, in close proximity to the *racR* promoter. In turn, this results in significant reduction in *racR* expression, which unblocks *ydaS* and *ydaT* expression. These genes' function is unknown, but we predict indirectly YdaT activity leads to cell toxicity and loss of their fitness and viability. Under physiological conditions, *ydaT* is completely silent and its transcript is undetectable. The genetic elements taking part in transcriptional cross-talking were identified and the expression of *racR – ydaST* operon was studied extensively by *in vivo* and *in vitro* approaches.

Our results demonstrate an apparent example of horizontal gene transfer leading to adventitious TF cross-talk with negative effects on the recipient's viability or even its death. More broadly, this study represents an experimentally-accessible model of a regulatory constraint on horizontal gene transfer.

1. Wstęp

1.1. Transkrypcja

Transkrypcja to pierwszy etap ekspresji genów. Jest to proces, w którym sekwencja DNA genu jest przepisywana na cząsteczki mRNA. Trzy kluczowe etapy transkrypcji to: inicjacja, elongacja i terminacja (Browning & Busby, 2004). Transkrypcja może być regulowana negatywnie – przez represory, jak i pozytywnie – przez aktywatory. Regulacja transkrypcji ma zasadnicze znaczenie dla zapewnienia ekspresji genów we właściwych ilościach i w odpowiednim czasie (Seshasayee et al., 2011). Głównym enzymem transkrypcyjnym jest polimeraza RNA (RNAP). Składa się ona z 5 podjednostek: dwóch podjednostek α , które rozpoznają czynniki regulatorowe; podjednostki β , która katalizuje syntezę RNA, podjednostki β' , która wiąże niespecyficznie DNA oraz podjednostki ω , która tworzy część rdzeniową enzymu (Cramer, 2002). Aparat transkrypcyjny *Escherichia coli* składa się z: rdzenia polimerazy RNA, siedmiu promotorów rozpoznających czynniki sigma (σ) oraz około trzystu czynników transkrypcyjnych.

1.2. Czynniki sigma (σ)

Czynniki σ determinują specyficzność promotora, czyli odcinka DNA położonego powyżej sekwencji kodującej genu i są integralną częścią maszyny transkrypcyjnej. Czynniki σ , który jest składnikiem holoenzymu, czyli kompletnego aktywnego enzymu, składa się z części białkowej, czyli apoenzymu i części niebiałkowej, czyli kofaktora. Pełni on istotną rolę w ekspresji genów. Odpowiada za rozpoznawanie promotora oraz rekrutację holoenzymu do promotora, której rdzenną część stanowią podjednostki $\alpha 1$, $\alpha 2$, β , β' oraz ω , a do utworzenia holoenzymu niezbędny jest tzw. czynnik sigma, który może odwracalnie przyłączać się (w trakcie inicjacji) lub oddysocjowywać od holoenzymu. Białka te są niezbędne do tworzenia otwartego kompleksu przez RNAP (Typas & Hengge, 2006). Zapewniają najbardziej podstawowy poziom kontroli ekspresji dużej liczby genów (Seshasayee et al., 2011).

Rodzinę czynników σ dzieli się na dwie filogenetyczne rodziny: σ^{70} i σ^{54} . W większości transkrypcji, które odbywają się w komórce pośredniczy główny czynnik σ , który należy do rodziny σ^{70} . Wiele z genomów bakteryjnych koduje również kilka alternatywnych czynników σ , które regulują bardzo specyficzne zestawy genów pod

wpływem różnych czynników stresowych. Te czynniki również należą do rodziny σ^{70} . Czynniki σ należące do tej rodziny samodzielnie przeprowadzają otwartą stabilizację kompleksu, natomiast czynniki σ należące do rodziny σ^{54} , wymagają dodatkowych aktywatorów (Burgess, 2001). Czynniki σ^{70} są najbardziej rozpowszechnione wśród bakterii i posiadają wiele podgrup tych czynników, natomiast czynniki σ^{54} znajdują się tylko w około 65% zsekwencjonowanych genomów bakteryjnych i są zwykle pojedyncze (Gruber & Gross, 2003). Przykładowo, w genomie *E. coli* K-12 MG1655 znajduje się siedmiu członków rodziny σ : σ^{70} (RpoD), σ^{32} (RpoH), σ^{38} (RpoS), σ^{24} (RpoE), σ^{28} (RpoF/FliA), σ^{19} (FecI), oraz σ^{54} (RpoN) (Burgess, 2001).

W komórce bakteryjnej nieustannie toczy się konkurencja o liczbę cząsteczek apoenzymu RNAP, czyli części białkowej polimerazy RNA. Wynik konkurencji determinuje ekspresję genów. Konkurencja i jej powodzenie zależy od: stężenia czynników σ (Holland & Rather, 2008), obecności białka odłączającego czynnik σ białka anty-sigma (Rsd u *E. coli*) (Jishage & Ishihama, 1998), obecności modulujących wtórnych przekaźników, które są drobnocząsteczkowe takie jak ppGpp (czterofosforan guanozyny), które biorą udział w reakcji stresowej (Jishage et al., 2002), obecności małych niekodujących RNA (np. 6S RNA) (Wassarman & Storz, 2000), obecności innych białek (np. H-NS) (Shin et al., 2005), oraz zdolności czynnika σ do rozpoznawania ewolucyjnie rozbieżnych miejsc promotora (Typas & Hengge, 2006). Aby zapewnić prawidłowy przebieg procesu transkrypcji niezbędne są tzw. czynniki transkrypcyjne.

1.3. Czynniki transkrypcyjne

Czynniki transkrypcyjne (TFs, ang. *transcription factors*) są białkami, które pomagają uaktywnić lub wyciszyć poszczególne geny. Czynniki te wiążą się z konkretną sekwencją DNA w pobliżu genów docelowych, modulując inicjację transkrypcji, utrudniając lub ułatwiając polimerazie RNA wiązanie się z promotorem. TFs mogą aktywować lub hamować transkrypcję w zależności od miejsca ich wiązania. Każdy TF reguluje pewien zestaw genów (Seshasayee et al., 2011). Szacuje się, że genom *E. coli* koduje około 270 różnych czynników transkrypcyjnych, co odpowiada 8% genów kodujących białka u *E. coli* (Babu & Teichmann, 2003). Natomiast według Ishihama, 2009 występuje około 10% białek wiążących DNA.

1.4. Globalne białka regulatorowe w transkrypcji

Globalne białka regulatorowe wpływają na ekspresję genów w skali całego genomu w odpowiedzi na przychodzące sygnały komórkowe. Ustalają one priorytet procesów zachodzących w komórce pozwalających na przetrwanie. W sytuacji stresu wywołanego przez głód będą dążyły do wykorzystania alternatywnych źródeł węgla ustalając to za najwyższy priorytet. Wśród wszystkich czynników transkrypcyjnych wiążących DNA globalni regulatorzy celują w nieproporcjonalnie dużą liczbę genów (Seshasayee et al., 2011).

Przykładem TF o zasięgu globalnym jest Lrp. Został on po raz pierwszy zidentyfikowany jako regulator transportu aminokwasów (Anderson et al., 1976). Jego aktywność jest modulowana przez leucynę (Lin et al., 1992). U *E. coli* reguluje geny zaangażowane w metabolizm, transport aminokwasów oraz niektóre funkcje niemetaboliczne: np. biosynteza pilusów (Seshasayee et al., 2011). Zidentyfikowano około 140 miejsc wiązania się Lrp w chromosomie (Cho, Barrett, et al., 2008). Wykazano również, że brak leucyny zwiększa liczbę tych miejsc 3-4 krotnie (Calvo & Matthews, 1994), co z kolei wskazuje na fakt, iż tak naprawdę, to niezdefiniowane miejsce wiązania, a bardziej warunki środowiskowe bakterii będą definiować liczbę miejsc wiązania się Lrp w chromosomie (Seshasayee et al., 2011). Lrp i leucyna mogą oddziaływać na trzy różne sposoby. Po pierwsze: niezależna odpowiedź, gdy leucyna nie ma wpływu na działanie Lrp, po drugie: uzgodniona odpowiedź, gdy leucyna wzmacnia działanie Lrp, po trzecie: wzajemna odpowiedź, w której leucyna antagonizuje działanie Lrp. Lrp najczęściej występuje w dwóch formach: oktamerycznej (Lrp8) i heksadekamerycznej (Lrp16). Wiązanie leucyny sprzyja postaci oktamerycznej (Chen et al., 2001). Lrp może również zginać i owijać DNA (McFarland et al., 2008). Homolog Lrp występujący u *Bacillus subtilis* pomaga w tworzeniu mostków DNA (Tapias et al., 2000). Z tych powodów Lrp zaliczono go do grupy NAP (ang. *nucleoid-associated proteins*), czyli białek związanych z nukleoidami (Seshasayee et al., 2011).

Innym przykładem czynnika posiadającego zasięg globalny jest Crp. Crp jest znane również pod nazwą białko CAP (ang. *catabolite gene activator protein* lub *cAMP receptor protein*). Crp jest najczęściej występującym czynnikiem transkrypcyjnym u *E. coli* (Gama-Castro et al., 2008). Crp jest związany z regulacją operonu laktozowego oraz jest aktywowany przez wiązanie się cyklicznego AMP (cAMP) w odpowiedzi na głód spowodowany brakiem glukozy lub innymi czynnikami stresowymi (Seshasayee et al.,

2011). Jego najważniejsza funkcja to udział kompleksu cAPM-CRP w represji katabolicznej, ale również w syntezie rzęski, produkcji enterotoksyn oraz wykorzystywaniu kwasów tłuszczowych (Zheng et al., 2004). Badania wiązania się Crp do genomu *E. coli* ujawniły około 70 takich miejsc (Grainger et al., 2005). Na podstawie zdolności CRP do zginania DNA zalicza się go również do grupy NAP (S. H. Lin & Lee, 2003).

Kolejnym przykładem jest regulator Fis (ang. *factor for inversion stimulation*). Jest to uniwersalne białko wiążące DNA, które może wpływać na wiele procesów, w tym na transkrypcję (Seshasayee et al., 2011). U *E. coli* jest głównym regulatorem zmian dotyczących wzrostu komórki (Schneider et al., 1997). Ekspresja tego czynnika jest zależna od fazy wzrostu i jest najwyższa podczas fazy logarytmicznej (Azam et al., 1999). Dla większości genów jest aktywatorem, jednak istnieją także takie, których ekspresja jest hamowana (Kahramanoglou et al., 2011). Fis negatywnie reguluje własny operon, hamując wiązanie polimerazy RNA. Przeprowadzono analizę, w której zidentyfikowano 894 regiony oddziaływania białka Fis w genomie *E. coli*. Ta analiza wskazała, że odcinki A i AT są ważnym sygnałem dla wiązania się białka Fis. Analiza transkryptomu *E. coli* wykazała, że delecja genu *fis* wpływa na około 21% genów (Cho, Knight, et al., 2008). Białko Fis ma zdolność do zginania DNA, w regionach które są potrzebne w procesach inwersji DNA oraz procesach aktywujących transkrypcję promotorów RNA (Kahramanoglou et al., 2011).

Kolejnym przykładem regulatora o zasięgu globalnym jest H-NS. Jest to globalny negatywny modulator ekspresji genów w rodzinie pałeczek jelitowych. Jest on obecny we wszystkich fazach wzrostu *E. coli* (Seshasayee et al., 2011). Wpływa na strukturę DNA, tworząc mostki DNA-HN-S-DNA i wzmacniając skręcenie helisy (Dorman, 2004). H-NS wykazuje powinowactwo do rejonów DNA bogatych w pary A/T (Atlung & Ingmer, 1997). Analiza genomu *E. coli* wykazała istnienie dwóch typów regulacji genów za pomocą tego białka: krótkie regiony wiążące zapewniają łagodny wpływ na ekspresję, podczas gdy długie regiony prowadzą do całkowitego wyciszenia transkrypcji (Kahramanoglou et al., 2011).

1.5. Miejscowi regulatorzy transkrypcji

W wielu przypadkach, globalne czynniki transkrypcyjne są wspomagane przez inne czynniki transkrypcyjne działające lokalnie tworząc sieć zależności. Są to bardzo

specyficzne, działające miejscowo czynniki, mające ograniczony zakres regulacyjny, który zwykle obejmuje kilka genów lub operonów. Lokalne czynniki transkrypcyjne stanowią większość repertuaru regulacyjnego w genomach bakteryjnych (Seshasayee et al., 2011).

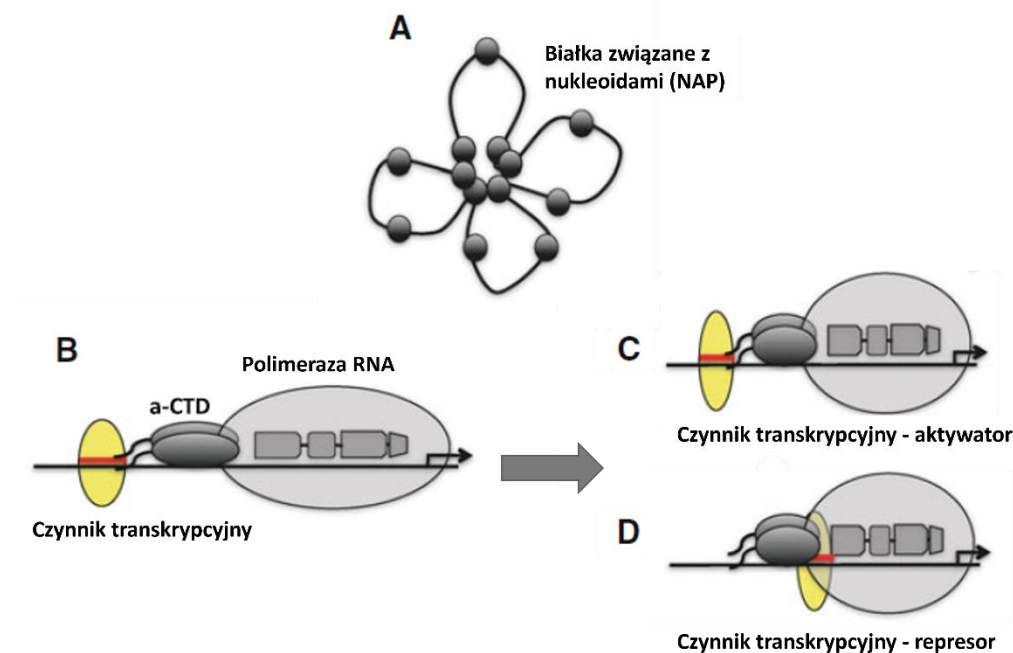
Jednym z najlepiej poznanych TF o zasięgu lokalnym jest LacI. Reguluje on ekspresję operonu *lac*. Operon *lac* w *E. coli* zawiera geny, które są zaangażowane w metabolizm laktozy. Operon ten ulega ekspresji tylko w obecności laktozy i jednoczesnym braku glukozy. Posiada on dwa regulatory: represor LacI i białko CAP. LacI działa jako czujnik glukozy. Blokuje transkrypcję operonu, ale działa jak represor, gdy nie jest obecna laktoza. Represor *lac* wyczuwa allolaktoze (Seshasayee et al., 2011).

Kolejnym, równie dobrze poznany czynnikiem o działaniu lokalnym jest TrpR. Reguluje on operon syntezy tryptofanu. TrpR wykrywa ilość wolnego tryptofanu. Operon *trp* jest aktywowany, gdy poziom tryptofanu jest niski, a hamowany, gdy jego poziom jest wysoki. Gdy poziom tryptofanu jest wysoki represor Trp wiąże się z tryptofanem (Seshasayee et al., 2011). Powstaje wtedy aktywny represor *trp*, który blokuje ekspresję operonu tryptofanowego.

Innym przykładem TF, który działa miejscowo jest OmpR. Jest to przykład miejscowego czynnika dwuskładnikowego systemu regulacji. We współpracy z EnvZ kontroluje osmoregulację oraz syntezę białka błony zewnętrznej. EnvZ jest kinazą histydynową zlokalizowaną w błonie wewnętrznej. Po aktywacji ulega autofosforylacji poprzez ATP, a następnie aktywuje OmpR. Fosforylowany OmpR ma bardzo wysokie powinowactwo do wiązania się z regionami regulatorowymi genów poryn *ompF* i *ompC*, dzięki czemu reguluje ich transkrypcję (Foo et al., 2015).

Jeszcze innym czynnikiem o działaniu miejscowym jest MarR. Jest to czynnik, który uczestniczy w kontrolowaniu ekspresji kilku genów zaangażowanych w nadawanie oporności na antybiotyki (Reyes-Fernández & Schuldiner, 2020), odpowiedź na stres oksydacyjny i stres wywołany obecnością jonów metali ciężkich. MarR jest częścią operonu *marRAB* i ujemnie reguluje swoją ekspresję (Aleksun & Levy, 1997).

Podsumowując, czynniki można podzielić w zależności od ich zasięgu działania (Ryc. 1.). Rycina 1.A. przedstawia białka związane z nukleoidami (NAP), które wiążą się z chromosomalnym DNA i pomagają w tworzeniu przestrzennej struktury DNA. Rycina 1.B przedstawia klasyczne czynniki transkrypcyjne, które pomagają polimerazie RNA w regulacji transkrypcji. Te klasyczne czynniki, można podzielić na aktywatory (Ryc. 1.C) lub represory (Ryc. 1.D).



Ryc. 1. Schemat przedstawiający podział TFs względem ich aktywności. **A** – NAP, **B** – TF, **C** – TF, który działa jak aktywator, oraz **D** – TF, który działa jako represor. a-CTD – podjednostka alfa (RpoA) inicjuje składanie RNAP poprzez dimeryzację, tworząc platformę, na której podjednostki beta mogą wchodzić w interakcje. Schemat opracowałam na podstawie Seshasayee et al., 2011.

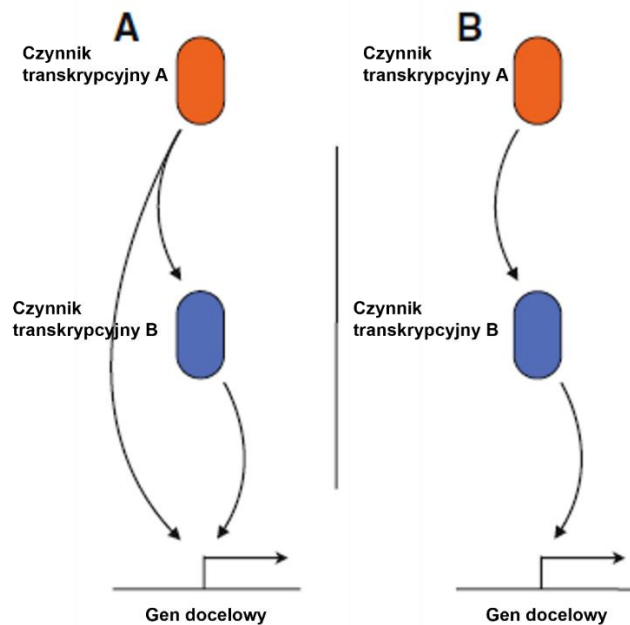
1.6. Bakteryjne sieci regulacji transkrypcji

Oddziaływanie czynników transkrypcyjnych z genami wyznacza profil ekspresyjny komórki bakteryjnej, a to przekłada się na jej fenotyp. Istnieje wiele algorytmów badających i identyfikujących motywy zależności pomiędzy czynnikami transkrypcyjnymi, a genami. Najlepiej scharakteryzowanym motywem jest motyw sieciowy (Seshasayee et al., 2011).

Motywy sieciowe to powtarzające się obwody oddziaływań regulacyjnych między czynnikami transkrypcyjnymi, a docelowymi genami (Shen-Orr et al., 2002). Takie motywy zostały pierwotnie określone w *E. coli*, w których wykrywano je jako wzorce połączeń. Te połączenia występują częściej niż się oczekuje w sieciach losowych (Seshasayee et al., 2011).

Jednym z najważniejszych motywów jest pętla sprzężenia w przód (wyprzedzająca) (FFL, ang. *Feed Forward loop*), w której TF A reguluje TF B, a A i B regulują docelowy gen C (Ryc. 2. A.). Druga sieć regulacji tzw. kaskada pojawia się, gdy TF A reguluje TF B, który z kolei reguluje gen docelowy (Ryc. 2. B.). TF najwyższego poziomu – A jest regulatorem globalnym (Seshasayee et al., 2011). Usunięcie z takiej

sieci czynnika globalnego doprowadza do utraty wielu pętli typu FFL (Shen-Orr et al., 2002).



Ryc. 2. Schemat regulacji sieciowych genów *E. coli*, w których wyróżnia się dwie podsieci. Panel **A** - pętla sprzężenia w przód (FFL) i panel **B** - kaskada. Schemat opracowałam na podstawie Seshasayee et al., 2011.

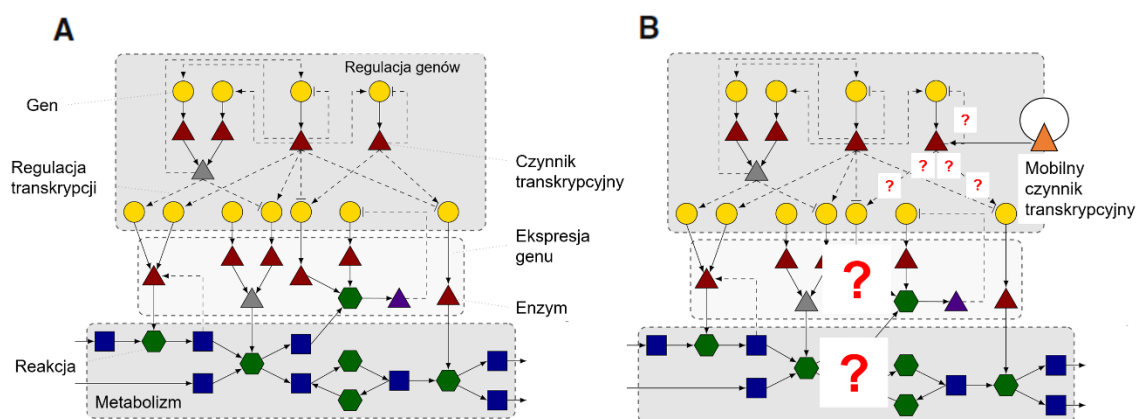
Przedstawione motywy sieciowe łączą się ze sobą tworząc całe systemy regulacji. W przypadku *E. coli* poszczególne sieci regulacji oddziałują ze sobą tworząc warstwę gęsto połączonych czynników transkrypcyjnych. Taki układ określa się jako motyw wielowejściowy (ang. *multi-input motifs* – MiMs). W takiej sytuacji każdy TF reguluje wiele genów docelowych, a każdy cel jest kontrolowany przez wiele czynników transkrypcyjnych. U *E. coli* jest kilka odrębnych MiMs-ów z setkami genów wyjściowych, z których każdy odpowiada za ważną funkcję biologiczną np. reakcję na stres. Długie kaskady regulacyjne są rzadkością w przypadku *E. coli*. Prawdopodobnie warunkowane jest to koniecznością szybkiej adaptacji komórek bakteryjnych do zmiennych warunków środowiska (Seshasayee et al., 2011).

Istotną kwestią w postrzeganiu bakteryjnej regulacji genów jest to, że polimerazy RNA w komórce *E. coli* jest bardzo mało. Szacuje się, że około 1500 do około 11 500 cząsteczek polimerazy przypada na komórkę w zależności od warunków wzrostu. Badania molekularne i biofizyczne w ciągu ostatnich 50 lat wyjaśniły różne mechanizmy modulowania ekspresji genów: niektóre mechanizmy pozwalają na precyzyjne dostrojenie poziomów, inne w ogóle nie pozwalają się modulować, z uwagi na ważność

procesu. Będą one miały restrykcyjne przejścia pomiędzy aktywnymi, a nieaktywnymi stanami transkrypcji (Seshasayee et al., 2011).

1.7. Zakłócenie bakteryjnych sieci regulacji transkrypcji

Na poziomie genu mutacja jest czynnikiem wpływającym na selekcję naturalną. Niektóre z powstających mutacji będą zapewniały bakterii przetrwanie, inne będą dla niej letalne. Jednak na poziomie sieci regulacji transkrypcji ujawnia się zmienność w selekcji naturalnej poprzez nieuporządkowaną aktywność regulatorową w postaci reakcji krzyżowej, tzw. "cross-talk". Cross-talk można rozumieć jako zakłócenie (dezorganizację, destabilizację) ustalonej sieci regulatorowej. Jest to sytuacja, w której jeden TF wpływa na pętlę regulatorową innego czynnika transkrypcyjnego, zaburzając regulację w tej pętli. Może to być przypadkowa reakcja krzyżowa czynnika transkrypcyjnego z jego niepowiązanym (poza docelowym) miejscem w DNA, który ma wpływ na zmianę w transkrypcji genów nie docelowych (ang. *non-target* lub *off-target*) (Ryc. 3.). Wydaje się, że tzw. „cross-talk” pomiędzy sieciami regulatorowymi jest rzadki, ponieważ nieprawidłowa regulacja genów najczęściej zmniejsza szanse na przeżycie w środowisku, do którego organizm już jest przystosowany (Taylor et al., 2022). Istnieją także mechanizmy wspomagające izolacje pomiędzy sieciami regulatorowymi. Jednym z takich przykładów jest występowanie specyficznych reszt białkowych, które określają powinowactwo wiązania. Dotyczy to dwuskładnikowych systemów transdukcji sygnału. Przedstawione wyniki wskazują, że preferencyjne powinowactwo wiązania jest podstawą wierności wiązania w układach dwuskładnikowych (Willett et al., 2013). Innym przykładem mechanizmów izolujących sieci jest wprowadzanie mutacji adaptacyjnych. Mutacje takie mają na celu izolacje pomiędzy sieciami w celu rozprzestrzenienia się rodzin paralogicznych, czyli homologicznych (o wspólnym pochodzeniu ewolucyjnym) pomiędzy białkami sygnalizacyjnymi (Capra et al., 2012). Cross-talk może nadać gospodarzowi nowe cechy przystosowawcze, które mogą ułatwić ewolucję genowej sieci regulatorowej (Wagner, 2021). W konsekwencji występuje równowaga pomiędzy utrzymaniem specyficzności wiązania czynników transkrypcyjnych, a ich pokrewnymi miejscami wiązania. Powoduje to wysoką wydajność działania sieci regulatorowej oraz rozluźnienie specyficzności, aby umożliwić zmiany oraz adaptację w tej sieci (Taylor et al., 2022). Równowaga natomiast jest tutaj zjawiskiem umownym i jest zależna od preferencji i korzyści komórki bakteryjnej.



Ryc. 3. A – Schematyczne przedstawienie biologicznych procesów zaangażowanych w integrację metaboliczno-regulacyjną w sieci *E. coli*. Procesy regulacji genów obejmują przede wszystkim geny (żółte kółka), białka lub czynniki transkrypcyjne (brązowe trójkąty) oraz kompleksy (szare trójkąty), natomiast procesy metaboliczne są definiowane przez małe cząsteczki (niebieskie kwadraty) i katalizujące reakcje biochemiczne (zielone sześciokąty). Interakcje między procesami regulacyjnymi i metabolicznymi można scharakteryzować głównie za pomocą białek (fioletowe trójkąty) pełniących funkcję enzymów i regulatorów. Łąca regulacyjne są reprezentowane jako linie przerywane, łąca kodujące i związane z reakcją są pokazane jako linie ciągłe. **B** – Wprowadzenie do sieci regulacji mobilnego czynnika transkrypcyjnego za pomocą HTG. Jeśli wejście obcego TF (pomarańczowy trójkąt) przeszkadza jednemu lub więcej TF gospodarza mogą wystąpić negatywne konsekwencje dla komórki biorcy. Konsekwencje te mogą mieć różny wpływ na wszystkie przedstawione procesy. Mogą również prowadzić do śmierci gospodarza. Schemat opracowałam na podstawie Grimbs et al., 2019.

Sieci regulatorowe genów ewoluują poprzez ciągłe zmiany (McAdams et al., 2004). Zmiana w połączeniach (ang. *rewiring*), czyli specyficzne połączenie pomiędzy elementami regulatorowymi w obrębie jednej sieci regulacji jest tak samo podatne na adaptację jak organizacja genomu (Iglér et al., 2018). Jednak czynniki, które determinują potencjał ewolucyjny do zmiany połączeń genetycznych w sieciach regulatorowych - czyli pojawienia się nowych interakcji regulatorowych w obrębie tej samej sieci regulatorowej są słabo poznane. Ze względu na powszechność zjawiska horyzontalnego transferu genów (HTG) istnieje bardzo duże podobieństwo w sekwencji i strukturze w obrębie białek występujących w różnych sieciach regulatorowych (Taylor et al., 2022). Proces ten skutkuje wieloma regulatorami kontrolującymi nakładające się regulony (Huang et al., 2019). Zjawisko to naturalnie zapewnia uniwersalność czynnika oraz jego „niewierność” (ang. *promiscuity*). „Niewierne” wiązanie (ang. *promiscuous binding*) występuje wtedy, gdy TF wiąże się z miejscem poza docelowym – niespecyficznym (ang. *off-target*). Występowanie zjawiska „niewiernego” wiązania zapewnia szybką ewolucję, pozwalając na nowe powiązania regulatorowe w sieciach regulatorowych (ang. *rewiring*). Jednym ze skutków tzw. cross-talk jest uzyskanie przez TF nowej funkcji regulatorowej,

na którą może być akurat selekcja w środowisku występowania bakterii (Taylor et al., 2022). Jeżeli nowo powstałe połączenie jest korzystne dla gospodarza, selekcja może oddziaływać na powstanie mutacji, aby zwiększyć powinowactwo do nowego miejsca (Lamrabet et al., 2019). Sekwencje promotorowe mogą ewoluować poprzez bardzo wiele zmian mutacyjnych (Yona et al., 2018). Homologia strukturalna pomiędzy czynnikami transkrypcyjnymi umożliwia modyfikowanie ich domen wiążących, aby umożliwić im zdobycie nowych celów molekularnych (Rosanova et al., 2017). Tworzenie potencjalnych miejsc wiązania czynników transkrypcyjnych jest bardzo powszechne, a korzystne połączenia mogą zostać wzmocnione podczas presji selekcyjnej (Taylor et al., 2015).

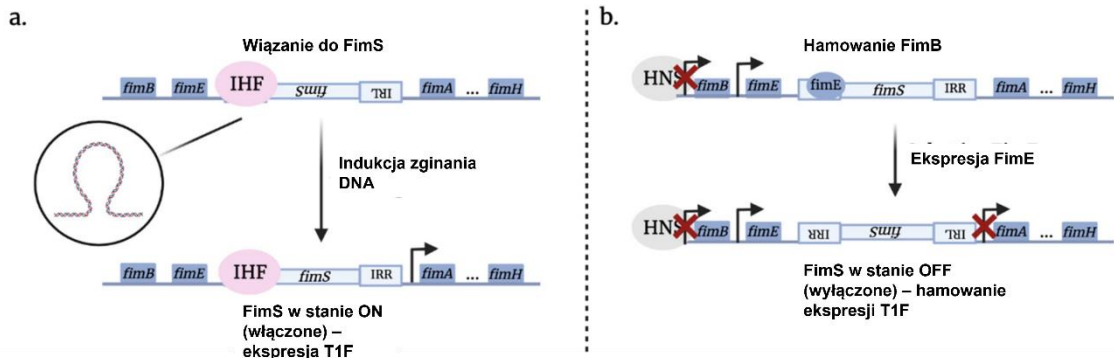
Czynniki transkrypcyjne wiążą się z sekwencją DNA poprzez domenę wiążącą (Perez-Rueda et al., 2018). Jednak miejsca wiązania nie mają jednolitej sekwencji, tylko są odmianami sekwencji konsensusowych (Kotelnikova et al., 2005) o różnym powinowactwie wiązania (Harari et al., 2010). W szczególności sekwencje konsensusowe bogate w pary A/T wykazują niższą specyficzność wiązania, a zatem bardziej prawdopodobne jest wiązanie się wielu czynników do tych sekwencji (Kurafeiski et al., 2019). Sytuacje w których dochodzi do wiązania się wielu czynników do jednej sekwencji, powodują zakłócenia w globalnych sieciach regulacji transkrypcji, na skutek konkurencji o wiązanie się do DNA.

1.8. Zakłócenie sieci globalnych regulatorów transkrypcji

Globalni regulatorzy transkrypcji charakteryzują się szerokim spektrum działania – modułują wiele procesów zachodzących w komórce. Są to białka wielozadaniowe i celują w dużą liczbę genów. Zakłócenie działania globalnego regulatora będzie niosło za sobą szereg zaburzeń w sieciach regulatorowych.

Przykładem zakłócenia działania globalnego regulatora jest zmienność fazowa fimbrii typu 1 u *E. coli*. Badania wykazały, że Lrp wiąże się z wysokim powinowactwem do dwóch miejsc w obrębie przełącznika *fim* (miejsca *fim1* i *fim2*), a wiązanie się z tymi miejscami stymuluje rekombinację. W badaniach *in vivo* pokazano, że leucyna stymuluje zmianę *fim*, to jednak powoduje nieznaczny spadek wiązania Lrp do zmiany miejsc *fim*. W badaniach wykazano również wiązanie się Lrp z trzecim regionem sąsiadującym z miejscami *fim1* i *fim2*, co hamuje rekombinację poprzez zakłócenie wiązania Lrp z tym nowo scharakteryzowanym regionem, jednocześnie wywierając niewielki lub żaden

wpływ na wiązanie Lrp z miejscami *fim1* i *fim2* (Roesch & Blomfield, 1998). Wykazano również wpływ IHF oraz H-NS, innych globalnych regulatorów na ekspresję fimbrii typu 1 (Ryc. 4.) (Bessaiah et al., 2021).



Ryc. 4. Mechanizm działania regulatorów stresu na ekspresję fimbrii typu 1. **A** – Regulatory mogą wiązać się bezpośrednio do *fimS*, aby wpływać na zmianę fazy. Na przykład IHF wiąże się z regionami *fimS*, aby wywołać zginanie DNA, które ułatwia rekombinację przełącznika *fim*. W środowisku pozbawionym składników odżywczych, indukowana jest ekspresja IHF, zwiększając zmienność faz i wpływając na ekspresję fimbrii typu 1. **B** – Regulatory mogą blokować ekspresję rekombinazy, aby wpływać na zmianę fazy *fimS*. W przypadku H-NS regulator może wiązać się z promotorem *FimB*, aby zablokować jego ekspresję, co skutkuje większą produkcją *FimE*. Ponieważ *FimE* ułatwia zmianę fazy *fimS* w stanie wyłączonym (OFF) - fimbrie typu 1 zostaną zahamowane. Schemat opracowałam na podstawie Bessaiah et al., 2021.

Innym przykładem jest regulacja transportu kwasów tłuszczowych u *E. coli*. Obejmuje ona trzy odrębne systemy regulatorowe, wyspecjalizowany system FadR i dwa systemy globalne: ArcA i CRP-cAMP. Działanie FadR jest analogiczne do LacI, czyli jest klasycznym represorem i wycisza ekspresję *fadD* i *fadL*. FadR pełni dwie funkcje regulacyjne w metabolizmie kwasów tłuszczowych. Hamuje system degradacji kwasów tłuszczowych (*fad*) i aktywuje szlak syntezy nienasyconych kwasów tłuszczowych. Regulacja cAMP polega na wykorzystaniu najprostszego źródła węgla, głównie glukozy. W jej obecności (lub pod nieobecność CRP) pozostałe dwa regulatory mają niewielką zdolność regulowania transkrypcji poprzez reakcję krzyżową (ang. *cross-talk*). Zaobserwowano wzajemne oddziaływanie między FadR, a systemem ArcA-ArcB, co skutkowało zróżnicowaną regulacją niektórych genów β -oksydacji regulonu *fad*. Promotory zarówno genów *fadL*, jak i *fadD* (ich produkty białkowe są wymagane do transportu długołańcuchowych kwasów tłuszczowych) zawierają dwa potwierdzone eksperymentalnie miejsca wiązania FadR oraz miejsca wiązania dla ArcA i CRP-cAMP. Pomimo obecności podwójnych miejsc wiązania FadR tylko w niewielkim stopniu reguluje ekspresję tych genów, co wskazuje, że liczba miejsc wiązania nie determinuje siły regulacyjnej. Zarówno badania *in vitro* i *in vivo* wskazują, iż kompleks CRP-cAMP

bezpośrednio aktywuje ekspresję *fadL* i *fadD*, jak również genu β -oksydacji, *fadH* (Feng & Cronan, 2012).

Inny TF – Fis, może wiązać się z miejscem rozpoznawanym przez IHF. Działa to również w drugą stronę – IHF może wiązać się z miejscem konsensusowym Fis. Najciekawsze jest to, że to samo dotyczy IHF, który również był zdolny do rozpoznawania miejsc wiązania Fis (Monteiro et al., 2019).

Innym przykładem jest białko Cnu. Jest to białko nukleoidowe, które wiąże się z chromatyną i reguluje ekspresję genów wirulencji *E. coli* w odpowiedzi na zmiany temperatury lub siły jonowej. Cnu składa się z trzech helis α . Wraz ze wzrostem siły jonowej helikalna struktura Cnu ulega destabilizacji, zwłaszcza na końcach helis. Dominujące miejsca wiązania H-NS, zlokalizowane są w helisie 3 (Bae et al., 2008).

1.9. Zakłócenie sieci miejscowych regulatorów transkrypcji

Reakcję krzyżową (ang. *cross-talk*) opisano również dla układów dwuskładnikowych: CpxA-CpxR i EnvZ-OmpR. Wiele bakterii posiada dwuskładnikowe systemy sygnalizacyjne, które składają się z par kinaza histydynowa-regulator odpowiedzi. Wysoki poziom podobieństwa sekwencji między niektórymi systemami stwarza możliwość niepożądanego przenikania się kinazy histydynowej z niespokrewnionym regulatorem. Zbadano interferencję tych dwóch niespokrewnionych czynników: CpxA, a OmpR. Badania pokazują, że istnieją mechanizmy, które mają za zadanie hamować wzajemne oddziaływanie między tymi dwoma białkami, pomimo tego występuje pomiędzy nimi reakcja krzyżowa (ang. *cross-talk*) (Siryaporn & Goulian, 2008).

Innym przykładem reakcji krzyżowej (ang. *cross-talk*) z udziałem TF działającego lokalnie jest regulon *mar-sox-rob* w *E. coli*. Czynniki transkrypcyjne MarA, SoxS i Rob odgrywają rolę w koordynowaniu wielu mechanizmów przetrwania w odpowiedzi na różnego rodzaju stres. Znane są trzy oddzielne systemu regulujące ekspresję genów MarA, SoxS i Rob. Jednak wiele badań wykazało, że nie działają one w izolacji, ale są współregulowane poprzez reakcję krzyżową (ang. *cross-talk*) na poziomie transkrypcyjnym (Chubiz et al., 2012).

Co ciekawe, nie znalazłam przykładów zaburzenia sieci regulatorowych dotyczących LacI oraz TrpR.

1.10. Różne „oblicza” zjawiska reakcji krzyżowej (ang. *cross-talk*)

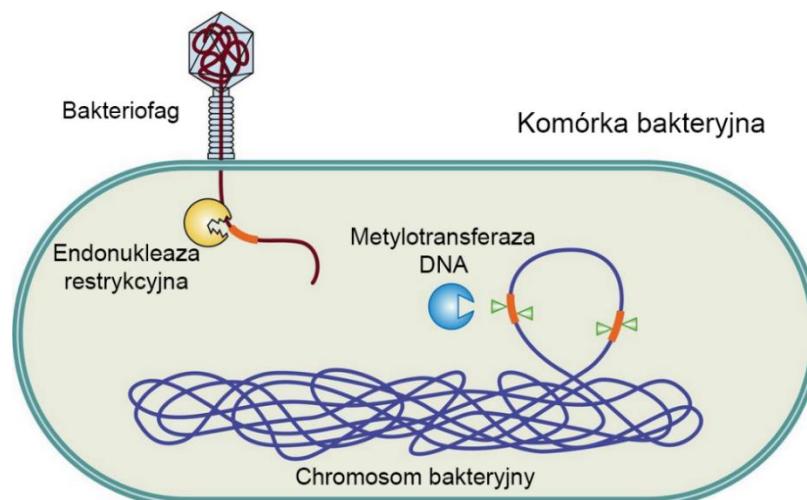
E. coli może być zarówno komensalem, jak i patogenem. Różnorodność tej bakterii wynika z nabywania i ekspresji wielu genów, w tym związanych z wirulencją. Regulacja ekspresji w tym aktywowania lub hamowania tych genów jest bardzo złożona i zależna od wielu czynników. Regulację transkrypcji traktuje się jako globalną, kiedy regulator oddziałuje na dużą liczbę genów, albo lokalną, kiedy regulator kontroluje ekspresję tylko określonego zestawu genów, często we własnym operonie (Holden & Gally, 2004).

Zjawisko reakcji krzyżowej (ang. *cross-talk*) może mieć różne „oblicza”. Przykładowo, wykazano pozytywny efekt pomiędzy klastrami genów adhezyn. Wiele operonów fimbrii P i S (i innych) może występować u patogenów *E. coli*, aby poradzić sobie z różnorodnością receptorów, a wprowadzenie ich „uniwersalności” ułatwiłoby przetrwanie. Ta reakcja krzyżowa oznaczałaby również, że mogą istnieć „główne” operony adhezyne, które są odpowiednio kontrolowane przez sygnały środowiskowe. Wykazano również negatywny efekt działania tych zjawisk. Kluczową rolą negatywnego jest to, aby różne typy adhezyn nie ulegały ekspresji w tym samym czasie przez tę samą komórkę bakteryjną, ponieważ może to zwiększyć szanse na wywołanie odpowiedzi immunologicznej, a także, że jeden z nich może zahamować drugi. Taka sytuacja na pewno wpłynęłaby negatywnie na żywotność komórki bakteryjnej (Holden & Gally, 2004). Transkrypcja i regulacja genów muszą zachodzić nieustannie w komórce bakteryjnej, aby zapewnić jej przetrwanie i adaptację do szybko zmieniających się warunków środowiska (Seshasayee et al., 2011). Jednym ze sposobów na regulację przepływu genów jest obecność systemów restrykcyjno-modyfikacyjnych. Pełnią one kluczową rolę w horyzontalnym transferze genów jako modulatory tego zjawiska, dozując napływ obcego DNA.

1.11. Funkcje systemów R-M

Systemy restrykcyjno-modyfikacyjne (R-M) są wszechobecne u prokariotów i często uważane za ich uproszczone układy odpornościowe. Systemy R-M są klasyfikowane na cztery główne typy w oparciu o skład ich podjednostek, rozpoznawanie sekwencji, pozycję cięcia, wymagania dotyczące kofaktorów i specyficzność substratową, jednak Typ II jest najlepiej poznany. Do tej pory poznanych jest prawie 4000 enzymów o około 300 różnych specyficznościach (Roberts et al., 2010). Zaobserwowano,

że około 90% genomów zawiera co najmniej jeden system R-M, a około 80% zawiera wiele systemów R-M (<http://rebase.neb.com/rebase/rebase.html>). Różnorodność i powszechność występowania w całym królestwie prokariotów wskazuje na ich sukces jako mechanizm obronny przed inwazją obcego DNA. Jednym z osiągnięć w ewolucji i różnorodności drobnoustrojów jest zdolność bakterii do rozróżnienia obcego DNA od ich własnego. Bakterie wykształciły strategie ograniczania stałego napływu obcego DNA do ich komórki. Postuluje się, że główną funkcją systemów R-M jest ochrona genomu gospodarza przed infekcją bakteriofagową. Systemy te odgrywają ważną rolę w horyzontalnym transferze genów (HTG) wpływając na napływ obcego DNA, pozytywnie i negatywnie. Na funkcję relatywnie najprostszego spośród systemów R-M, Typu II, składa się aktywność dwóch niezależnych enzymów: endonukleazy restrykcyjnej (REazy) oraz ochronnej metylotransferazy DNA (MTazy). Oba enzymy rozpoznają tę samą specyficzną sekwencję DNA. REaza przecina niezmetylowaną sekwencję DNA, zaś MTaza chroni materiał genetyczny gospodarza przed degradacją przez pokrewną REazę, poprzez przyłączanie do niego grup metylowych (Vasu & Nagaraja, 2013) (Ryc. 5.).



Ryc. 5. Zasada działania systemu R-M w komórce bakteryjnej. Systemy R-M są systemami obronnymi. Rozpoznają one stan metylacji obcego DNA (np. bakteriofagowego). Zmetylowane sekwencje są rozpoznawane jako własne, podczas gdy sekwencje obce są cięte przez REazę. Schemat opracowałam na podstawie Vasu & Nagaraja, 2013.

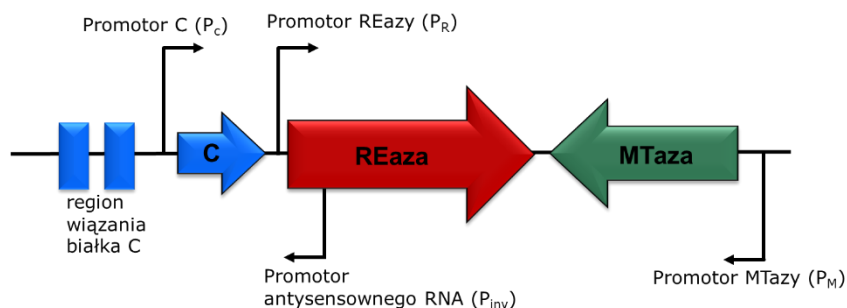
1.12. Systemy R-M Typu II

Na funkcję systemu R-M Typu II składa się aktywność dwóch niezależnych enzymów: REazy i MTazy. Generalnie REazy Typu II są homodimeryczne lub homotetrameryczne i przecinają nici DNA w miejscu docelowym lub w jego pobliżu

(Vasu & Nagaraja, 2013). Enzymy Typu II rozpoznają ściśle zdefiniowane, najczęściej palindromowe sekwencje DNA (Roberts et al., 2003). Są one najszerzej zbadane i najczęściej stosowane w inżynierii genetycznej (Vasu & Nagaraja, 2013). Szczególnym przykładem systemów restrykcyjno-modyfikacyjnych Typu II są te posiadające dodatkowy komponent: wyspecjalizowane białko regulatorowe, które może regulować aktywność REazy, bądź też MTazy (Rezulak et al., 2016). Zachwianie równowagi pomiędzy aktywnością REazy i MTazy może spowodować śmierć gospodarza, dlatego ekspresja genów systemów R-M musi podlegać ściślejszej regulacji, często poprzez właśnie dedykowany regulator (Mruk & Kobayashi, 2014).

1.13. Organizacja genetyczna systemu R-M Csp231I

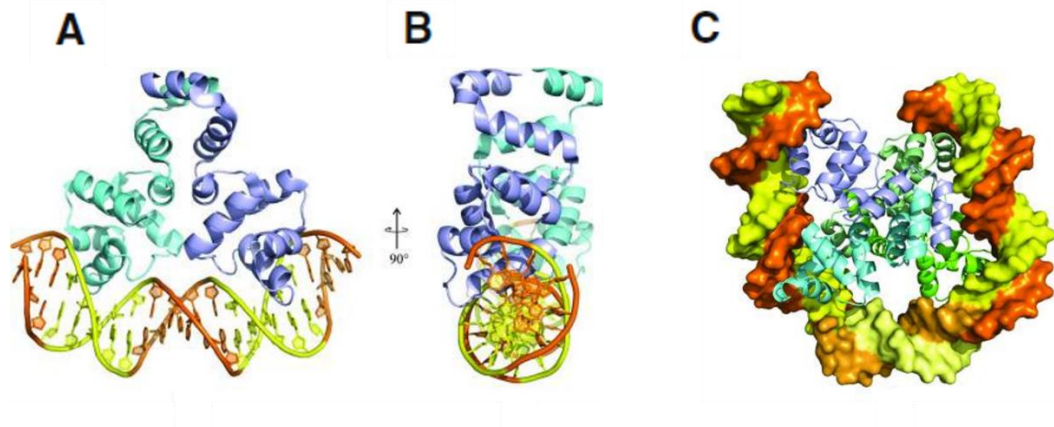
Naszym modelowym systemem R-M tego typu jest system Csp231I, który odkryto w chromosomie bakterii *Citrobacter sp.* RFL 231 (Mruk & Kaczorowski, 2007). W systemie tym oba enzymy rozpoznają sekwencję 5'-AAGCTT-3'. System R-M Typu II Csp231I jest zbudowany z trzech genów. Zbieżnie ułożonego genu białka regulatorowego C (ang. *controller*) wraz z genem endonukleazy restrykcyjnej oraz przeciwnie ułożonego genu metylotransferazy DNA (Ryc. 6.).



Ryc. 6. Organizacja genetyczna systemu restrykcyjno-modyfikacyjnego Csp231I z *Citrobacter sp.* RFL231. Gen białka C poprzedzony jest regionem wiązania tego białka. Geny REazy oraz MTazy są zorientowane zbieżnie. Oznaczyłam również część promotorów. Schemat opracowałam na podstawie Rezulak et al., 2016.

Systemy R-M muszą podlegać ściślejszej kontroli. Zakłócenie tej kontroli może prowadzić do śmierci komórki bakteryjnej. Regulacja ta musi być już obecna na etapie tzw. „instalacji” systemu R-M w nowym gospodarzu. Koniecznym jest, aby ekspresja MTazy była najwcześniej, aby zmetylować całkowicie genom gospodarza, zanim dojdzie do ekspresji REazy. Dopiero przy wysokim poziomie metylacji, produkcja REazy jest bezpieczna. Opóźnienie aktywności restrykcji jest kluczowe dla przeżycia gospodarza.

Następnie stężenie MTazy powinno się zmniejszyć, aby zapobiec metylowaniu obcego DNA. Oba te enzymy funkcjonują analogicznie do układu toksyna – antytoksyna, gdyż nadmierna aktywność REazy lub niedostateczna aktywność MTazy może doprowadzić do śmierci komórki gospodarza (Mruk & Kobayashi, 2014). To właśnie czynniki transkrypcyjne poprzez regulację genów mają za zadanie utrzymanie balansu np. pomiędzy toksyną i antytoksyną w systemach T-A lub R-M. Jednym z trzech mechanizmów kontroli produkcji enzymów systemów R-M jest czynnik transkrypcyjny C. System R-M Csp231I posiada taki dodatkowy gen, który koduje czynnik transkrypcyjny C. Białko C jest wyspecjalizowanym białkiem regulatorowym, które kontroluje poziom ekspresji genów systemu R-M. Regulator C jest represorem własnej transkrypcji. Białko C wiąże się do sekwencji C-box znajdującej się w rejonie własnego promotora jako tetramer (Rezulak et al., 2016). Jego strukturę przedstawiłam na Rycinie 7. Gen C znajduje się przed genem REazy, z którym jest wspólnie transkrybowany (Vijesurier et al., 2000). Jest autoregulatorem własnego genu i pośrednio genu REazy. Jest to dedykowany czynnik transkrypcyjny działający miejscowo w obrębie systemu R-M. Nie jest to globalny czynnik transkrypcyjny (Rezulak et al., 2016). Białko C wiąże się do sekwencji konsensusowej: 5'-CTAAG-N5-CTTAG-3'(Shevtsov et al., 2015).



Ryc. 7. Struktura kompleksu C.Csp231I – DNA. **A** – Dimer białka związany z dupleksem DNA. **B** – Dimer (a) obrócony o 90°. **C** – Oddziaływanie dimer-dimer i model tetrameryczny. Przy wiązaniu tetrameru tworzy się pętla DNA w obrębie tzw. sekwencji rozdzielającej dwa odwrócone powtórzenia (ang. *spacer*). Rycinę opracowałam na podstawie Shevtsov et al., 2015.

2. Cel pracy

Celem rozprawy doktorskiej było zbadanie mechanizmu molekularnego odpowiedzialnego za proces formowania się wydłużonych komórek bakteryjnych (tzw. filamentów) w szczepie *Escherichia coli* MG1655 niosących geny systemu restrykcyjno-modyfikacyjnego Typu II Csp231I, pochodzącego ze spokrewnionej bakterii *Citrobacter* sp. RFL231.

Obejmowało to realizację następujących zadań szczegółowych:

- ✓ identyfikację elementów genetycznych odpowiedzialnych za proces filamentacji komórek *Escherichia coli*;
- ✓ zbadanie wpływu filamentacji na żywotność komórek gospodarza;
- ✓ charakterystykę mechanizmu molekularnego prowadzącego do filamentacji w warunkach *in vivo* oraz *in vitro*.

3. Materiały

3.1. Szczepy bakteryjne

Tabela 1. Szczepy bakteryjne.

Szczepy bakteryjne	Charakterystyka	Źródło
<i>Escherichia coli</i>		
MG1655	K-12 F ⁻ λ ⁻ <i>ilvG</i> ⁻ <i>rfb-50 rph-1</i>	(Jensen, 1993)
MG1655Δ <i>rac</i>	Jak MG1655WT, ale z delecją całego regionu faga <i>rac</i> (210 p.z.)	(Guo et al., 2014)
DH5α	F ⁻ λ ⁻ <i>endA1 thi-1 recA1 relA1 gyrA96 glnV44 deoR nupG φ80dlacZΔM15 Δ(lacZYA-argF) U196 hsdR17</i> (rK ⁻ mK ⁺)	(Woodcock et al., 1989)
MP060	Δ <i>attHK022::</i> (P _{sulA} - <i>yfp</i>)	(Pleška et al., 2016)
MP064	Δ <i>attHK022::</i> (P _{sulA} - <i>yfp</i>), Δ <i>recA</i>	(Pleška et al., 2016)
BL21(DE3)	F ⁻ <i>ompT gal dcm lon hsdS_B(r_B⁻ m_B⁻) λ</i> (DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) [<i>malB</i> ⁺] _{K-12} (λ ^S)	Novagen
ER2566	F ⁻ λ ⁻ <i>fhuA2 [lon] ompT lacZ::t7 gene1 gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr-73::miniTn10-Tet^S)2 R(zgb-210::Tn10--Tet^S) endA1 [dcm]</i>	Novagen
Rosetta(DE3)	F ⁻ , <i>ompT, hsdS_B</i> (r _B ⁻ m _B ⁻), <i>gal, dcm</i> (DE3) pRARE	Novagen
MM294	<i>glnV44(AS), rfbC1, endA1, spoT1, thi-1, hsdR17, creC510</i>	(Bachmann, 1972)
MC1061	<i>araD139 Δ(ara, leu)7697, ΔlacX74, galU, galK, hsdR, strA</i>	(Casadaban & Cohen, 1980)

3.2. Plazmidy bakteryjne

Tabela 2. Plazmidy bakteryjne.

Nazwy plazmidu	Charakterystyka	Źródło
pBR322	Replikon pMB1, Amp ^R , Tet ^R	(Bolivar et al., 1992)
p18	Jak pBR322, Tet ^R , ale geny systemu R-M Csp231I WT	(Mruk & Kaczorowski, 2007)
p23	Jak p18, ale gen C zmutowany jako S(16)→A; Q(17)→A; E(18)→A. Białko C niezdolne do wiązania się z DNA	(Rezulak et al., 2016)
p24	Jak p18, ale gen REazy został przerwany (R ⁻ M ⁺ C ⁺)	(Rezulak et al., 2016)

p30	Jak p18, ale ΔC i ΔC -box	(Rezulak et al., 2016)
p32	Jak p18, ale z wyłączonym promotorem P_{R1} genu <i>R</i>	(Rezulak et al., 2016)
p34	Jak p30, ale z wyłączonym promotorem P_{R1} genu <i>R</i>	Monika Rezulak
pBAD-33	Wektor z promotorem indukowalnym arabinozą P_{BAD} , ori pACYC, Cm^R	(Guzman et al., 1995)
pBAD-CWT	Jak pBAD-33, ale gen <i>C</i> pod promotorem P_{BAD}	(Rezulak et al., 2016)
pBAD-Cmut	Jak pBAD-CWT, ale gen <i>C</i> zmutowany jako S(16)→A; Q(17)→A; E(18)→A. Białko <i>C</i> niezdolne do wiązania się z DNA	(Rezulak et al., 2016)
pBAD-RacR	Jak pBAD-33, ale gen <i>racR</i> pod promotorem P_{BAD}	(Negri et al., 2019)
pBAD-RacRmut	Jak pBAD-RacR, ale przerwany gen <i>racR</i> , trawiono BamHI, następnie uzupełniono końce fragmentem Klenow i połączono w procesie ligacji	Ewa Wons
pBAD-YdaS	Jak pBAD-33, ale gen <i>ydaS</i> pod promotorem P_{BAD}	Ta praca
pBAD-YdaSmut	Jak pBAD-YdaS, ale gen <i>ydaS</i> zmutowano przy użyciu <i>ydaSkRFwd</i> i <i>ydaSkRrev</i> , K(37)→A; R(40)→A	Ta praca
pLEX3B	Wektor do testowania fuzji translacyjnych i transkrypcyjnych z genem reporterowym <i>lacZ</i> , pBR322 ori, Amp^R	(Diederich et al., 1994)
pLEX3BRacRI-long	Jak pLEX3B, ale z genami <i>racR-ydaS-T</i> (420 p.z.), które zostały amplifikowane przy użyciu PFracRpLEX1 oraz PRracRpLEX3 i wprowadzone poprzez trawienie enzymami restrykcyjnymi XhoI i HindIII, gen <i>racR</i> w fuzji z genem reporterowym <i>lacZ</i>	Ta praca
pLex3Bbox1MUT	Jak pLex3BRacRI-long, ale ze zmianą w sekwencji box1 (CTTAG → CGCAT) Mutageneza przy użyciu PFracRframeIpLEX1 oraz PRracRframeIpLEX2 wprowadzając miejsce restrycyjne SphI.	Ta praca
pLex3BracRydaS	Jak pLEX3B, ale z genami <i>racR</i> i <i>ydaS</i> (704 p.z.), które zostały amplifikowane przy użyciu PFracRXhoI oraz PRydaSPstI i wprowadzone poprzez trawienie enzymami restrykcyjnymi XhoI i PstI, gen <i>racR</i> w fuzji z genem reporterowym <i>lacZ</i>	Ta praca
pLex3BracR(inakty)ydaS	Jak pLex3BracRydaS, ale przerwany gen <i>racR</i> , trawiono BamHI, następnie uzupełniono końce fragmentem Klenow i połączono w procesie ligacji	Ta praca
pLex3BracRydaST	Jak pLEX3B, ale z genami <i>racR</i> , <i>ydaS</i> i <i>ydaT</i> (984 p.z.), które zostały amplifikowane przy PFracRXhoI oraz PRydaTPstI i wprowadzone poprzez trawienie enzymami restrykcyjnymi XhoI i PstI, gen <i>racR</i> w fuzji z genem reporterowym <i>lacZ</i>	Ta praca

pLex3BracR(inakty)ydaST	Jak pLex3BracRydaST, ale przerwany gen <i>racR</i> , trawiono BamHI, następnie uzupełniono końce fragmentem Klenow i połączono w procesie ligacji	Ta praca
pLex3BracR(inakty)ydaS(inakty)ydaT	Jak pLex3BracR(inakty)ydaST, ale gen <i>ydaS</i> przerwany, P(29)→A;P(30)→A;	Ta praca
pET28a(+)	Wektor z indukowanym IPTG promotorem T7, pBR ori, kan ^R	Novagen
pET28a(+)RacR	Jak pET28a(+) z wklonowanym genem <i>racR</i> z sekwencją znacznika histydynowego (His-6), który został amplifikowany przy użyciu PFRacRpetNcoI i PRracRpetEcoRI i wprowadzony poprzez trawienie enzymami restrykcyjnymi NcoI i EcoRI	Ta praca
pDSW360	Gen <i>ftsZ</i> w fuzji z białkiem reporterowym GFP, pBR ori, kan ^R	Samuel J. Craven

3.3. Oligonukleotydy i substraty DNA

Tabela 3. Oligonukleotydy i substraty DNA wykorzystane w pracy

Eksperyment	Nazwa oligonukleotydu	Sekwencja (5'→3')
Oligonuklotydy użyte do EMSA	PRydaSPstI	GCCTCTGCAGTGCTGATTGCCAC CGACAAC
	PFRacRfluo1Cy5 znacznik	Cy5-CCGCCTTTGATTTGACGGATC
	RACRplexC	GCGACGCCTAACGCAATTGTCAA TAGGCTG
	PFunspecFLUO3-Cy5	Cy5-TTCACTGCAATATTTTCTCGG
	PRunspecFLUO4	AAAACATTCATTACTGACCGC
	PFRacRpLEX1	AAATCTCGAGTTGCTTGAATGAA TAGTTCTC

Substraty EMSA	PRydaSPstI PFracRfluo1Cy5znacznik	<i>racR</i> promotor/operator/rejon intergenowy (IGR) i część sekwencji kodującej <i>racR</i> (271 p.z.) CATTGCTGATTGCCACCGACAA CAGCACAAGCTTGCTTGAATGAA TAGTTCTCTTTTTTCATCGAATGA ACTCCAAAAACACACAGAAATAT TAGGCGACGCCTAACGCAATTGT CAATAGGCTGTGCCTAATGCAGT AAGGGTAGGGATTGCCTAATGTA ATGCGCATAGGAGAATATTAAGC AATGCTTAGTGGTAAAGACTTAG GCCGAGCGATAGAGCAGGCCATT AACAAAAAATCGCATCGGGATC CGTCAAATCAAAGGCGG
	PFunspecFLUO3-Cy5 PRunspecFLUO4	Sekwencja niespecyficzna <i>racR</i> - (209 p.z.) AAAACATTCATTACTGACCGCAT GATTAAAGTCTTCCTCGGTCGTGA TGGTTTACCTGTAAAGGCGGAGT CATGGTGATGACTAAGAAAATAA AATGTGCTTACCACCTTTGCAA AAAGACGTTGAAGAAAGCAAAG CTATTGAAAGAATGCTTCACTTCA TGCACGGGATTTTATCAAAGAC GAACCGAGAAAATATTGCAGTGA A
Oligonukleotydy użyte do metody śladu DNA (ang. <i>footprinting</i>)	PRydaSPstI	GCCTCTGCAGTGCTGATTGCCACCGACAAC
	PFracRfluo1Cy5 znacznik	Cyanine5-CCGCCTTTGATTTGACGGATC
	RACRplexC	GCGACGCCTAACGCAATTGTCAATAGGCTG
Substraty DNA użyte do metody śladu DNA (ang. <i>footprinting</i>)	PRydaSPstI PFracRfluo1Cy5znacznik	<i>racR</i> promotor/operator/rejon intergenowy (IGR) i część sekwencji kodującej <i>racR</i> (271 p.z.) CATTGCTGATTGCCACCGACAA CAGCACAAGCTTGCTTGAATGAA TAGTTCTCTTTTTTCATCGAATGA ACTCCAAAAACACACAGAAATAT TAGGCGACGCCTAACGCAATTGT CAATAGGCTGTGCCTAATGCAGT

			AAGGGTAGGGATTGCCTAATGTA ATGCGCATAGGAGAATATTAAGC AATGCTTAGTGGTAAAGACTTAG GCCGAGCGATAGAGCAGGCCATT AACAAAAAATCGCATCGGGATC CGTCAAATCAAAGGCGG
		RACRplexC PFracRfluo1Cy5znacznik	<i>racR</i> sekwencja kodująca (175 p.z.) GCGACGCCTAACGCAATTGTCAA TAGGCTGTGCCTAATGCAGTAAG GGTAGGGATTGCCTAATGTAATG CGCATAGGAGAATATTAAGCAAT GCTTAGTGGTAAAGACTTAGGCC GAGCGATAGAGCAGGCCATTAAC AAAAAATCGCATCGGGATCCGT CAAATCAAAGGCGG
Klonowania	pET- RacR	PFracRpetNcoI	GAGCTACCATGGTTAGTGGTAAA GACTTAGGCCG
		PRracRpetEcoRI	GCTCGAATTCCTTAGTGGTGGTGGT GGTGGTGAGTTCCAGTTTTTGTGA TGTTTC
	pBAD- YdaS	PR-pBADydaS1	AAAGAGCTCGAATCATTTCCTC TTGATGTTTCAGG
		PF-pBADydaS2	GGCTCTAGAGGAGTTCATTCGAT GAAAAAAGAGAAC
	pLEX3B RacRI- long	PFracRpLEX1	AAATCTCGAGTTGCTTGAATGAA TAGTTCTC
		PRracRpLEX3	CCTAAAGCTTTTTTGTATCTGAAT TGGTGG
	pLex3Bra cRydaS	PFracRXhoI	GAATCCTCGAGTTTTCCGTCAGAC CATATAAG
		PRydaSPstI	GCCTCTGCAGTGCTGATTGCCAC CGACAAC
	pLex3bra cRydaST	PFracRXhoI	GAATCCTCGAGTTTTCCGTCAGAC CATATAAG
		PRydaTPstI	CCCACTGCAGCCCTTTTTCGGCTG CTAGGGC

Mutagenезy	pBAD-YdaSmut	ydaSkRFwd	GTAGCGCAATTGCCTGCCGAGAG A
		ydaSkRrev	CCCCGCGATCCATTGATTACGCT TGG
	pLex3Bb ox1MUT	PFracRframeIpLE X1	GAATATTAAGCAATGCGCATGCG TAAAGACTTAGGCCG
		PRracRframeIpLE X2	CGGCCTAAGTCTTTACGCATGCG CATTGCTTAATATTC

3.4. Wzorce masowe

- Wzorzec masowy DNA przedstawiający prążki w zakresie 250-10000 p.z. – Perfect Plus 1 kb DNA Ladder (Eurx)
- Wzorzec masowy białek przedstawiający prążki w zakresie 10-180 kDa – PageRuler Prestained Protein Ladder (Thermo Scientific)

3.5. Antybiotyki

Ampicylina (Sigma) używana w stężeniu 100 µg/ml
 Chloramfenikol (Sigma) używany w stężeniu 35 µg/ml
 Tetracyklina (Sigma) używana w stężeniu 12,5 µg/ml
 Kanamycyna (Sigma) używana w stężeniu 50 µg/ml

3.6. Podłoża wzrostowe

3.6.1. Pożywka płynna LB (Luria Broth), pH 7,2

Trypton	10 g
NaCl	10 g
Ekstrakt drożdżowy	5 g
Woda destylowana	do 1000 ml

Pożywkę po doprowadzeniu pH do wartości 7,0 jałowiono w autoklawie przez 30 min przy nadciśnieniu 1 atm.

3.6.2. Pożywka stała LA (Luria Agar), pH 7,2

Trypton	10 g
NaCl	10 g
Ekstrakt drożdżowy	5 g
Agar bakteryjny	15 g
Woda destylowana	do 1000 ml

Pożywkę po doprowadzeniu pH do wartości 7,0 jałowiono w autoklawie przez 30 min przy nadciśnieniu 1 atm.

3.6.3. Pożywka minimalna M9 (skład na 100 ml)

5% casamino acids	2 ml
1 M MgSO ₄	100 µl
0,1 mM CaCl ₂	100 µl
1 mg/ml tiamina	200 µl
1× sole	20 ml
50 % glicerol	800 µl
woda jałowa	77 ml

3.7. Bufory i żele do elektroforez

3.7.1. Elektroforeza DNA w żelu agarozowym

3.7.1.1. Bufor elektrodowy TBE

Roztwór 10 × stężony	Roztwór 1 × stężony
Tris 108 g (890 mM) Kwas borowy 55 g (889 mM) 0,5 M EDTA pH 8.0 40 ml (0,02 mM) Woda destylowana do 1000 ml	Bufor 10× TBE 100 ml Woda destylowana 900 ml

3.7.1.2. Żel agarozowy 0,8%

Agaroz	0,8 g (0,8%)
Bufor TBE 1×	100 ml
Bromek etydyny	0,5 µg/ml, czyli 5 mg/ml

3.7.1.3. Barwnik obciążający do elektroforezy agarozowej 6 × stężony

Sacharoza	40 g 40 %
Błękit bromofenolowy	0,001 g 0,001 %
EDTA pH 8,0	200 mM

3.7.2. Elektroforeza białek w żelu poliakrylamidowym

3.7.2.1. Bufor SDS-PAGE pH 8,3

Roztwór 10 × stężony	Roztwór 1 × stężony
SDS 10 g (35 mM) Tris 30 g (25 mM) Glicyna 144 g (192 mM) Woda destylowana do 1000 ml	Bufor 10× SDS-PAGE 100 ml Woda destylowana 900 ml

3.7.2.2. Bufor Tris-HCl pH 8,8 (dolny)

Tris	181,6 g (1,5 M)
SDS	4 g (0,4%)
Woda destylowana	do 1000 ml

3.7.2.3. Bufor Tris-HCl pH 6,8 (górnny)

Tris	60,6 g (0,5 M)
SDS	4 g (0,4%)
Woda destylowana	do 1000 ml

3.7.2.4. Roztwór 30% akrylamidów

Akrylamid	29 g
N,N-metyleno-bis-akrylamid	1 g
Woda destylowana	do 1000 ml

3.7.2.5. Żel poliakrylamidowy rozdzielający (dolny) 12,5%

Bufor Tris-HCl pH 8,8	5,6 ml
Roztwór 30% akrylamidów	9,7 ml
Woda destylowana	7,1 ml
TEMED	90 μ l
10% nadsiarczan amonu (APS)	180 μ l

3.7.2.6. Żel poliakrylamidowy zagęszczający (górnny) 4%

Bufor Tris-HCl pH 6,8	2,5 ml
Roztwór 30% akrylamidów	1,4 ml
Woda destylowana	5,7 ml
TEMED	40 μ l
10% APS	80 μ l

3.7.2.7. Barwnik Laemmli 4 \times stężony

Tris-HCl pH 6,8	0,25 M
SDS	8 %
β -merkaptoetanol	8 %
100% Glicerol	40 %
Bromofenol blue	0,02 %
Woda destylowana	do 10 ml

3.7.2.8. Barwienie żeli białkowych

3.7.2.9. Barwnik Coomassie Brilliant Blue (CBB)

Roztwór 10 × stężony	Roztwór 1 × stężony
Coomassie Brilliant Blue 2,5 g Kwas octowy 100 ml Metanol 450 ml Woda destylowana do 1000 ml	Barwnik 10× CBB 500 ml Kwas octowy 200 ml Woda destylowana do 1000 ml

3.7.2.10. Odbarwiacz

Kwas octowy	100 ml
Metanol	200 ml
Woda destylowana	800 ml

3.7.3. Elektroforeza białek w żelu tricinowy

3.7.3.1. Bufor elektrodowy Tris-HCl pH 8,9 (dolny)

Tris	24,22 g (0,2 M)
SDS	0,3 g (0,3 %)
Woda destylowana	do 1000 ml

3.7.3.2. Bufor elektrodowy Tris (górny)

Tris	12 g (0,1 M)
Tricine	18 g (0,1 M)
SDS	1 g (0,1 %)
Woda destylowana	do 1000 ml

3.7.3.3. Bufor tricinowy pH 8,45

Tris	46,6 g (0,38 M)
SDS	0,3 g (0,3 %)
Woda destylowana	do 1000 ml

3.7.3.4. Żel tricinowy rozdzielający (dolny) 12,5 %

Bufor tricinowy pH 8,45	1,65 ml
Roztwór 30% akrylamidów	2 ml
Woda destylowana	0,8 ml
Glicerol	0,5 ml
TEMED	10 µl
10% APS	20 µl

3.7.3.5. Żel tricinowy zagęszczający (górny) 4 %

Bufor tricinowy pH 8,45	0,60 ml
Roztwór 30 % akrylamidów	1,25 ml
Woda destylowana	3,125 ml
TEMED	7,5 µl
10 % APS	12,5 µl

3.8. Bufory i roztwory używane do analizy białek

3.8.1. Bufor 2 x NPi pH 8,0

NaH ₂ PO ₄	0,05 M
NaCl	0,3 M
Imidazol	0,001 M
Woda destylowana	do 1000 ml

3.8.2. Bufor NPi-10

Bufor 2 x NPi pH 8,0	500 ml
Triton X-100	0,1 %
Glicerol	10 %
Imidazol (przed użyciem)	0,001 M
Woda destylowana	do 1000 ml

3.8.3. Bufor NPi-20

Bufor 2 x NPi pH 8,0	500 ml
Triton X-100	0,1 %
Glicerol	10 %
Imidazol (przed użyciem)	0,002 M
Woda destylowana	do 1000 ml

3.8.4. Bufor NPi-150

Bufor 2 x NPi pH 8,0	500 ml
Triton X-100	0,1 %
Glicerol	10 %
Imidazol (przed użyciem)	0,15 mM
Woda destylowana	do 1000 ml

4. Metody

4.1. Amplifikacja fragmentów DNA metodą PCR (reakcja łańcuchowa polimerazy ang. *Polymerase Chain Reaction*)

Skład mieszaniny reakcyjnej o całkowitej objętości 50 μ l:
plazmidowy DNA matrycowy 2 μ l
10 x bufor do polimerazy 10 μ l
dNTPs (2,5 mM) 5 μ l
starter 1 1 μ l
starter 2 1 μ l
polimeraza DNA (0,1 u/ μ l) 0,5 μ l
jałowa woda dejonizowana 30,5 μ l

Etapy reakcji PCR:

1. Denaturacja wstępna (3 min., 96°C)
 2. Denaturacja właściwa (30 sek., 96°C)
 3. Przyłączanie starterów (1 min, temperatura optymalna dla danej pary oligonukleotydów)
 4. Wydłużanie (przyłączanie polimerazy do kompleksu starter-matryca; czas zależny od długości amplifikowanego fragmentu, średnio 1 min/1000 p.z., 72°C, 30 cykli)
 5. Wydłużanie końcowe (3 min, 72°C)
- Po zakończonej reakcji PCR, próbę przechowywałam w zamrażarce (-20°C).

4.2. Mutageneza miejscowo-specyficzna

Mutageneza miejscowo-specyficzna umożliwia wprowadzenie zmiany nukleotydów w określonym miejscu znanej sekwencji DNA. Zastosowanie odpowiednio zaprojektowanych starterów umożliwiło mi wprowadzenie zmian w określonym miejscu matrycowego DNA, poprzez amplifikację całego plazmidu. Następnie, w celu usunięcia z mieszaniny niewielkiej ilości zmetylowanego, matrycowego DNA, przeprowadzałam trawienie z użyciem enzymu restrykcyjnego DpnI. Do 50 μ l mieszaniny reakcyjnej po PCR dodawałam 1 μ l enzymu DpnI i inkubowałam w 37°C przez 30 min. Następnie, oczyszczałam DNA z wykorzystaniem zestawu Clean-up PCR. Przed rozdziałem elektroforetycznym do próbki dodawałam barwnik obciążający. Rozdział elektroforetyczny prowadziłam w celu określenia wielkości produktu uzyskanego w wyniku mutagenezy miejscowo-specyficznej. Prążek o odpowiedniej wielkości wycinałam z żelu. DNA izolowałam z wykorzystaniem zestawu Gel-out. Odpowiedni szczep *E. coli* transformowałam oczyszczonym DNA. Następnie, wysiewałam całą mieszaninę transformacyjną na podłoże LA z dodatkiem odpowiedniego antybiotyku.

4.3. Elektroforeza agarozowa DNA

Rozdział elektroforetyczny DNA przeprowadzałam w 0,8% żelu agarozowym, w aparacie poziomym w buforze 1 x TBE. 0,54 g agarozy rozpuszczałam przez zagotowanie w 60 ml buforu TBE 1 × stężonego, a następnie po schłodzeniu i dodaniu bromku etydyny tak, aby końcowe stężenie wynosiło 0,5 µg/ml, żel wylewałam na płytkę aparatu. Próbki z barwnikiem obciążającym nanosiłam w studzienki żelu. Elektroforezę prowadziłam przy stałym natężeniu prądu o wartości 40 mA.

4.4. Oczyszczanie fragmentu DNA z żelu agarozowego

Interesujące mnie rozdzielone w żelu agarozowym fragmenty DNA wycinałam i oczyszczałam zestawem Gel-out (A&A Biotechnology) wg zaleceń producenta.

4.5. Izolacja DNA z żelu metodą elektroelucji

Interesujące mnie fragmenty rozdzielonego DNA umieszczałam w worku dializacyjnym z roztworem 1 x TBE (45,5 mM Tris, kwas borowy o pH = 8,0, 1 mM EDTA). Końce worka zamykałam klamrami i umieszczałam go w aparacie do elektroforezy agarozowej. Elektroelucję prowadziłam przy napięciu 50 V do momentu całkowitego wyjścia DNA z żelu. Następnie bufor TBE zawierający DNA przenosiłam do jałowych probówek i dodawałam 1/2 objętości mieszaniny roztworu 1:1 fenol-chloroform oraz 3 M octan sodu w ilości 1/10 wyjściowej objętości mieszaniny. Fazy rozdzielałam przez wirowanie (15 min., 12000 obr./min.). Fazę wodną przenosiłam do nowej probówki, następnie dodawałam dwie objętości mieszaniny 96% etanolu. Probówkę inkubowałam przez noc w -20°C. Wytrącony DNA osadzałam przez wirowanie (45 min., 12000 obr./min., 4 °C), po czym usuwałam supernatant i osad przepłukiwałam 70% etanolem. Plazmidowy DNA suszyłam przez 30 min., a następnie zawieszałam w 20 µl jałowej wody.

4.6. Trawienie DNA enzymami restrykcyjnymi

W skład mieszaniny reakcyjnej o końcowej objętości 20 µl wchodzi:

- woda destylowana
- roztwór DNA
- bufor o optymalnym składzie dla danego enzymu
- enzym restrykcyjny

Trawienie DNA enzymami restrykcyjnymi prowadziłam w optymalnej temperaturze dla danego enzymu restrykcyjnego. Czas trawienia zależny był od ilości DNA w próbce oraz aktywności enzymu restrykcyjnego.

4.7. Defosforylacja 5' końców DNA

Po trawieniu DNA plazmidowego enzymami restrykcyjnymi przeprowadzałam defosforylację końców 5' trawionych wektorów z wykorzystaniem alkalicznej fosfatazy CIP (EURx), zgodnie z zaleceniami producenta.

4.8. Ligacja fragmentów DNA trawionych enzymami restrykcyjnymi

Reakcje ligacji przeprowadzałam się w temperaturze 16°C przez noc w termocyklerze.

Skład mieszaniny reakcyjnej (10 µl):

10x bufor reakcyjny dla ligazy faga T4 1 µl

ligaza DNA faga T4 (10 u) 2 µl

insert 2 µl

wektor 5 µl

woda jałowa dejonizowana 1 µl

4.9. Przygotowanie komórek kompetentnych

W celu uzyskania komórek kompetentnych wybraną kolonię bakteryjną zaszczepiałam w 10 ml pożywki płynnej LB (Materiały 3.6.1.). Hodowlę inkubowałam przez noc w wytrząsarce rotacyjnej w temperaturze 37°C. Następnego dnia hodowlę nocną odmładzałam w 25 ml porcji pożywki LB przez zaszczepienie w proporcji 1:50 i umieszczałam ponownie w wytrząsarce rotacyjnej w temperaturze 37°C do uzyskania $OD_{600} = 0,4$. Po tym czasie schładzałam hodowlę w lodzie przez 15 min i przelewałam do jałowego falkonu, a następnie odwirowywałam w wirówce z chłodzeniem (4°C; 2000 × g; 10 min). Supernatant odrzucałam, a osad bakteryjny zawieszałam w 10 ml zimnego 0,1 M $CaCl_2$ i inkubowałam przez 30 min w lodzie. Zawiesinę bakterii odwirowywałam, zachowując te same warunki co poprzednio, a następnie osad bakteryjny zawieszałam w 3,5 ml 0,1 M $CaCl_2$ i 1,5 ml 50% glicerolu. Ostatecznie komórki kompetentne porcjowałam po 200 µl i przechowywałam w temperaturze -80°C.

4.10. Transformacja komórek kompetentnych

Do schłodzonej próbówki typu Eppendorf zawierającej 200 µl wybranej porcji szczepu komórek kompetentnych (Metoda 4.19.) dodawałam 1 µl (20 - 50 ng) wybranego

plazmidu. Mieszaninę inkubowałam w lodzie przez 30 min, a następnie poddawałam szokowi termicznemu w łaźni wodnej przez 3 min w temperaturze 42°C. Po tym czasie do komórek dodawałam 500 µl pożywki płynnej LB (Materiały 3.6.1.). Probówki z mieszaniną reakcyjną umieszczałam w termobloku ustawionym na temperaturę 37°C przez 45 min. Po tym czasie na płytce ze stałym podłożem LA (Materiały 3.6.2.) i odpowiednim antybiotykiem (Materiały 3.5.) za pomocą głaszczki wysiewałam 200 µl komórek bakteryjnych. Płytki inkubowałam przez noc w temperaturze 37°C.

4.11. Nadprodukcja białka w systemie bakteriofaga T7

Wybraną kolonię spośród otrzymanych transformantów zaszczepiałam w 10 ml pożywki płynnej LB (Materiały 3.6.1.) z dodatkiem odpowiedniego antybiotyku (Materiały 3.5.) i 0,4 % glukozy i wytrząsałam przez noc w wytrząsarce rotacyjnej w temperaturze 37 °C. Następnego dnia hodowlę nocną odmładzałam w 1000 ml pożywki LB przez zaszczepienie w proporcji 1:50 i umieszczałam ponownie w wytrząsarce rotacyjnej w temperaturze 37°C do uzyskania $OD_{600} = 0,4$. Po uzyskaniu pożądanej gęstości, hodowlę indukowałam 1 mM IPTG i ponownie prowadziłam inkubację przez 4 h. Po tym czasie hodowlę bakteryjną wirowałam w wirówce z chłodzeniem (4°C; 10000 × g; 20 min). Supernatant odrzucałam, a osad bakteryjny zawieszałam w buforze NPi-10 i przechowywałam w temperaturze – 80 °C do dalszych analiz.

4.12. Oczyszczanie białka

4.12.1. Liza komórek bakteryjnych

Lizę komórek bakterii przeprowadzałam za pomocą sonikacji przy użyciu urządzenia Sonicator Ultrasonic Processor XL2020 (Misonix). Osad bakteryjny rozmrażałam i zawieszałam w buforze NPi-10. Sonikację prowadziłam w lodzie przez 6 min, puls 10 sek., przerwa 1 min. Uzyskany roztwór wirowałam w wirówce z chłodzeniem (4°C; 11000 × g; 20 min), a uzyskany supernatant poddawałam dalszym analizom.

4.12.2. Chromatografia powinowactwa na złożu TALON

W celu zrównoważenia złoża TALON przemywałam je 10 objętościami jałowej wody, a następnie buforu NPi-10. Uzyskany lizat komórkowy dodawałam do probówki typu falkon zawierającego 1,5 ml złoża TALON i delikatnie zawieszałam, mieszając od czasu do czasu na lodzie (4°C; 20 min). Po tym czasie całość nanosiłam na kolumnkę 15 ml i

zbierałam przesącz. Następnie przemywałam złożę 10 objętościami buforu NPi-10 oraz NPi-20, a elucje białka prowadziłam przy 10 ml buforu NPi-150, gdzie frakcje zbierałam po 1 ml.

4.12.3. Dializa

Po przeprowadzonej analizie frakcji na żelu poliakrylamidowym wybierałam próbki, w których znajdowało się oczyszczane białko. Dializę nastawiałam w worku dializacyjnym na noc w buforze do dializ w temperaturze 4°C na mieszadle magnetycznym. Otrzymane białko porcjowałam do probówek typu Eppendorf po 500 µl i przechowywałam w 4°C do bieżących eksperymentów oraz w – 80 °C w celu przechowywania na dłuższy okres.

4.13. Pomiar stężenia białka

Ilość otrzymanego białka określałam poprzez zastosowanie krzywej wzorcowej zależności pomiędzy stężeniem molowym, a intensywnością prążka. Jako wzorzec wybrałam lizozym.

4.14. Sączenie molekularne

Stan oligomeryzacji białka sprawdzano przy użyciu chromatografii wykluczania (SEC). Sączenie molekularne przeprowadzono za pomocą urządzenia ÄKTA Pure 25 system (GE Healthcare) przy użyciu kolumny Superdex 75 10/300 GL (GE Healthcare). Kolumnę zrównoważono buforem o składzie: 50 mM K₂HPO₄/KH₂PO₄ pH 7,0, 150 mM NaCl. Następnie naniesiono na nią próbkę rekombinowanego białka RacR (monomer 18,5 kDa) o objętości 500 µL (0,5 mg). Rozdział prowadzono z szybkością przepływu 0,5 ml/min, eksperyment przeprowadzono w temperaturze 8° C. Czas retencji rekombinowanego białka monitorowano przy długości fali 280 nm. Kolumnę kalibrowano białkami o znanych masach cząsteczkowych: dehydrogenazą alkoholową (tetramer) 146,8 kDa; albumina surowicy bydlęcej 66 kDa; albumina jaja kurzego 43 kDa; inhibitor trypsyny 22 kDa i cytochrom C 12,4 kDa (Sigma-Aldrich, St. Louis, MO, USA). Urządzenie obsługiwał mgr Sebastian Dorawa.

4.15. Elektroforeza w żelu poliakrylamidowym w warunkach denaturujących

4.15.1. Przygotowanie żelu poliakrylamidowego

W pierwszej kolejności przygotowywałam 12,5% roztwór dolnego żelu poliakrylamidowego i wlewałam go między płyty aparatu do elektroforezy. Następnie górną warstwę pokrywałam niewielką ilością wody destylowanej, w celu pozbycia się pęcherzyków powietrza oraz zapewnienia odpowiednich warunków do polimeryzacji żelu. Po spolimeryzowaniu żelu dolnego usuwałam wodę i dodawałam przygotowany wcześniej 4% roztwór górnego żelu poliakrylamidowego, w którym umieszczałam przed jego polimeryzacją grzebień.

4.15.2. Wykonanie prób do elektroforezy białkowej

Aby przygotować próbki z nadprodukcji białka pobierałam 0,5 ml hodowli bakteryjnej przed indukcją oraz 1 ml po indukcji IPTG. Bakterie wirowałam (12 000 obr./min; 3 min), a uzyskany osad zawieszałam w 20 µl buforu Laemmli. Tak przygotowane próbki gotowałam w łaźni wodnej (100°C; 5 min), a następnie wirowałam (12000 obr./min; 30 s).

4.15.3. Elektroforeza białek w żelu poliakrylamidowym

Gotowe próbki nosiłam do studzienek przygotowanego wcześniej żelu poliakrylamidowego 10 µl przed indukcją oraz 5 µl po indukcji IPTG. W celu określenia wielkości białka na żelu, nosiłam 2 µl markera białkowego. Rozdział prowadziłam w roztworze 1× SDS-PAGE przy natężeniu prądu 180 mA, aż do wyjścia barwnika. Po rozdziale elektroforetycznym barwiłam żel w barwniku Coomassie Brilliant Blue, po czym umieszczałam go w odbarwiaczu, aż do odbarwienia się tła.

4.16. Barwienie fluorescencyjne LIVE/DEAD

SYTO®9 (Invitrogen, USA) o małej masie cząsteczkowej przenika do wnętrza komórki przez spójne błony cytoplazmatyczne, a PI o dużej masie cząsteczkowej przenika tylko do tych z uszkodzeniami w błonach. Zastosowanie zestawu LIVE/DEAD® umożliwia zróżnicowanie populacji bakterii na komórki żywe i komórki martwe pod względem różnic w ciągłości błon cytoplazmatycznych.

Nastawiałam hodowle nocne i indukowałam C (0,1 % ara) i GFP (1 mM IPTG) oraz uzupełniałam odpowiednimi antybiotykami. Następnego dnia wirowałam 1 ml hodowli przez 7 min 7 G temp. 22°C (po 2 próbki). Zawieszałam w 1 ml 0,85% NaCl. Wirowałam 1 ml hodowli przez 7 min 7 G temp. 22°C. Jedną zawieszałam w 70 % alkoholu izopropylowym (kontrola pozytywna – wszystkie martwe). Drugi osad zawieszałam z mastermixie – dodawałam 200 µl (mastermix: 997 µl 0,85% NaCl + 1,5 µl barwnika 1 + 1,5 µl barwnika 2). Inkubowałam 15 min w temperaturze pokojowej. Nakładałam 2 µl na szkiełko i oglądałam używając mikroskopu fluorescencyjnego DMB400B; Leica, Niemcy.

4.17. Przygotowanie preparatu do mikroskopii fluorescencyjnej

Przygotowywałam 100x rozcieńczony barwnik Cyto-9. Dodawałam 5 µl tego barwnika do 100 µl hodowli bakteryjnej. Następnie przygotowywałam 1% żel z agarozą. Na szkiełko nakrywkowe większe kładłam kostkę wykrojoną z żelu, następnie nakrapiałam 10 µl uprzednio przygotowanej hodowli z barwnikiem, następnie nakrywałam szkiełkiem nakrywkowym, a po nakropleniu olejku immersyjnego oglądałam preparat.

4.18. Indukcja ekspresji genów w hodowli bakteryjnej za pomocą arabinozy

Pojedynczą kolonię szczepu *E. coli* po transformacji z odpowiednim tłem genetycznym zaszczipiałam w 3 ml pożywki LB (Materiały 3.6.1.) suplementowanej odpowiednim antybiotykiem (Materiały 3.5.) oraz glukozą do stężenia 0,2%. Bakterie hodowałam przez noc w temp. 37°C. Hodowlę odmładzałam w świeżej pożywce LB w stosunku 1:50. Kontynuowałam hodowlę do OD₆₀₀=0,3. Następnie hodowlę rozdzielałam na odpowiednią ilość probówek z różnymi stężeniami arabinozy lub jej braku. Bakterie hodowałam 5 godzin, po tym czasie mierzyłam wartość OD₆₀₀ oraz przygotowywałam seryjne rozcieńczenia od 0 (brak rozcieńczenia) do 10⁻⁶ hodowli w 0,85% roztworze NaCl. Bakterie znajdujące się we wszystkich rozcieńczeniach nakrapiałam w ilości 10 µl na płytkę z podłożem LA (Materiały 3.6.2.) z odpowiednim antybiotykiem oraz arabinozą o końcowym stężeniu 0,1%.

4.19. Oznaczanie aktywności β-galaktozydazy (test ONPG)

Test ONPG opiera się na pomiarze enzymatycznego rozpadu substratu onitrofenylo-β-galaktopiranozydu (ONPG), co przejawia się zmianą zabarwienia roztworu z

bezbarnego na żółty (pomiar absorpcji przy długości fali 420 nm). Do testu przygotowywałam mieszaninę o składzie: 50 µl hodowli bakteryjnej (lub 50 µl pożywki minimalnej M9 w przypadku próby „ślepej”), 15 µl chloroformu, 7,5 µl 0,1% SDS i 400 µl buforu Z (60 mM Na₂HPO₄·2H₂O, 40 mM NaH₂PO₄·2H₂O, 10mM KCl, 1 mM MgSO₄·7H₂O, 50mM β-merkaptotanol). Próby wortexowałam i inkubowałam przez 10 min. w temperaturze 30°C. Następnie dodawałam po 50 µl roztworu ONPG do każdej próbki i obserwowałam zmianę zabarwienia próbek, odnotowując czas, w którym próbki zażółcały się. W celu zatrzymania reakcji dodawałam 150 µl 1 M Na₂CO₃. Następnie próbki wirowałam (3 min., 6000 obr./min.), po czym przenosiłam po 200 µl górnej fazy roztworu do płytki titracyjnej 96-dołkowej. Absorbancję mierzyłam przy długości fali 420 nm. Aktywności β-galaktozydazy obliczałam na podstawie wzoru:

$$\frac{(1000 * OD420)}{V * t * OD600} = \text{Jednostki Millera (J.M.)}$$

t- czas (minuty)

V-objętość hodowli (ml)

Poziom transkrypcji jest liniowo zależny od odpowiedniego wzrostu aktywności β-galaktozydazy. Modyfikowane Jednostki Millera (J.M.) uzyskiwałam poprzez pomiar w czasie aktywności β- galaktozydazy, wartością pomiaru był wynik równania R2.

4.20. Oznaczanie efektywności tworzenia kolonii bakteryjnych

Nocną hodowlę bakteryjną *E. coli* MG1655 (Materiały 3.1.) odmładzałam w 5 ml pożywki LB (Materiały 3.6.1.) uzupełnionej odpowiednim antybiotykiem (Materiały 3.5.) w stosunku 1:100. Następnie inkubowałam hodowlę z wytrząsaniem w temperaturze 37°C do uzyskania gęstości optycznej OD₆₀₀ = 0,2. Następnie przygotowywałam rozcieńczenia hodowli w roztworze soli fizjologicznej (0,85% NaCl). Z wybranych rozcieńczeń pobierałam 100 µl zawiesiny i całość wysiewałam na wysuszone płytki LA, uzupełnionym odpowiednim antybiotykiem (Materiały 3.5.). Płytki inkubowałam w 37°C przez noc. Uzyskane kolonie bakteryjne zliczałam i na podstawie otrzymanych wyników liczyłam wartość CFU (ang. *colony forming unit*) na podstawie wzoru:

$$\frac{N}{V * R} = CFU$$

N – ilość powstałych kolonii

V – objętość (ml)

R – rozcieńczenie

4.21. Test opóźnienia migracji DNA w żelu EMSA (ang. *Electrophoretic Mobility Shift Assay*)

Substratami DNA były dwuniciowe fragmenty amplifikowane za pomocą PCR, które były znakowane fluorescencyjnie z jednego końca znacznikiem Cy5 wprowadzonym przez startery PRydaSPstI i PFracRflu1cy5znacznik. Substrat WT (271 bp) obejmował cały promotor/operator/rejon między genowy *racR* (IGR) i część sekwencji kodującej *racR*. Reakcje zawierały 25 nM DNA. Białka CWT lub Cmut używałam w stężeniu 0, 264, 378, 422, 528 nM, a RacR używałam w stężeniu 0, 33, 231, 264, 300 nM. Wskazane stężenia białek zawieszałam w buforze wiążącym [10 mM Tris – HCl (pH 8,0), 50 mM NaCl, 10 mM MgCl₂, 1 µg poli (dIdC)] do końcowej objętości 20 µl i inkubowałam przez 20 min w 22°C. Próbkę poddawałam elektroforezie na 5% żelach poliakrylamidowych w buforze 0,5 x TBE w 22°C lub 1% żelach agarozowych. Wykrywanie DNA wyznakowanego Cy5 przeprowadzałam przy użyciu urządzenia do obrazowania Typhoon 9200 (Molecular Dynamics, USA).

4.22. Metoda śladu DNA (ang. *DNA footprinting*)

Reakcje wiązania białko-DNA przeprowadzałam w 20 µl w następującym buforze: [10 mM Tris – HCl (pH 8,0), 10 mM MgCl₂, 50 mM NaCl]. Jako matryce zastosowałam fragmenty PCR ze znacznikiem 5'Cy5 i zawierające sekwencję kodującą *racR* 175 p.z. (startery RACRplexC i PFracRFLUO1Cy5znacznik) lub sekwencję kodującą *racR* z IGR 271 p.z. (startery PRydaSPstI i PFracRFLUO1Cy5znacznik). Próbkę inkubowałam w 22°C przez 20 min, a następnie dodawałam 1 µl DNazy I (Eurx, Gdańsk, Polska) i inkubowałam dalej w 22°C przez 4 min. Reakcje kończyłam dodając EDTA do 25 mM i zatężałam przez odparowanie próżniowe. Następnie próbki ponownie zawieszałam w 40 µl buforu obciążającego (30 % formamid, 6 M mocznik, 10 mM NaOH), denaturowałam w 100° C przez 2 min i 7 µl nakładałam na 6 M mocznik 8% żel akrylamidowy wraz z reakcjami sekwencjonowania (zastosowałam startery wyznakowane 5'Cy-5, PFracRFLUO1Cy5znacznik, jako matryce zastosowałam plazmid pLEX3BracRlong, z terminatorem A (mieszanina dNTP zawierająca ddATP), terminatorem C (mieszanina dNTP zawierająca ddCTP), terminatorem G (mieszanina dNTP zawierająca ddGTP), T (mieszanina dNTP zawierająca ddTTP) z DNA Cycle Sequencing Kit, Jena Bioscience, Jena, Niemcy) i TERMIPol® DNA Polymerase firmy Solis BioDyne. Do skanowania i dokumentacji wykorzystywałam system obrazu Typhoon (GE Healthcare).

4.23. Relatywna restrykcja - oznaczanie efektywności tworzenia łyseinek bakteriofagowych

Nocną hodowlę bakteryjną *E. coli* MG1655 z plazmidem niosącym system R-M Csp231I odmładzałam w 5 ml pożywki LB uzupełnionej odpowiednim antybiotykiem w stosunku 1:100. Następnie inkubowałam hodowlę z wytrząsaniem w temperaturze 37°C do uzyskania gęstości optycznej $OD_{600} = 0,2$. W tym czasie przygotowywałam seryjne rozcieńczenia lizatu faga lambda (λ_{vir}) w buforze TM o składzie 10 mM Tris-HCl pH 7,4 oraz 10 mM $MgCl_2$. Z wybranych rozcieńczeń pobierałam 100 μ l zawiesiny bakteriofaga, a następnie dodawałam do 200 μ l hodowli. Następnie do probówek dodawałam 3 ml rozpuszczonego i schłodzonego agaru górnego (z połową dodanego agaru w pożywce LA) i całość przelewałam na wysuszone płytki LA. Płytki inkubowano w 37°C przez noc. Uzyskane łyseinki bakteriofagowe zliczałam i na podstawie otrzymanych wyników liczyłam wartość PFU (ang. *plaque forming unit*) na podstawie wzoru:

$$PFU = \frac{N}{V \times R}$$

gdzie:

N – liczba powstałych łyseinek

V – objętość lizatu fagowego (w ml)

R – rozcieńczenie lizatu fagowego

Wartość EOP obliczałam wg wzoru:

$$EOP = \frac{PFUpX}{PFUpBR32}$$

pX = badany konstrukt kontrolny bez systemu R-M

Wartość relatywnej restrykcji obliczałam wg wzoru:

$$RR = \frac{CFUp18(WT)}{CFUpX}$$

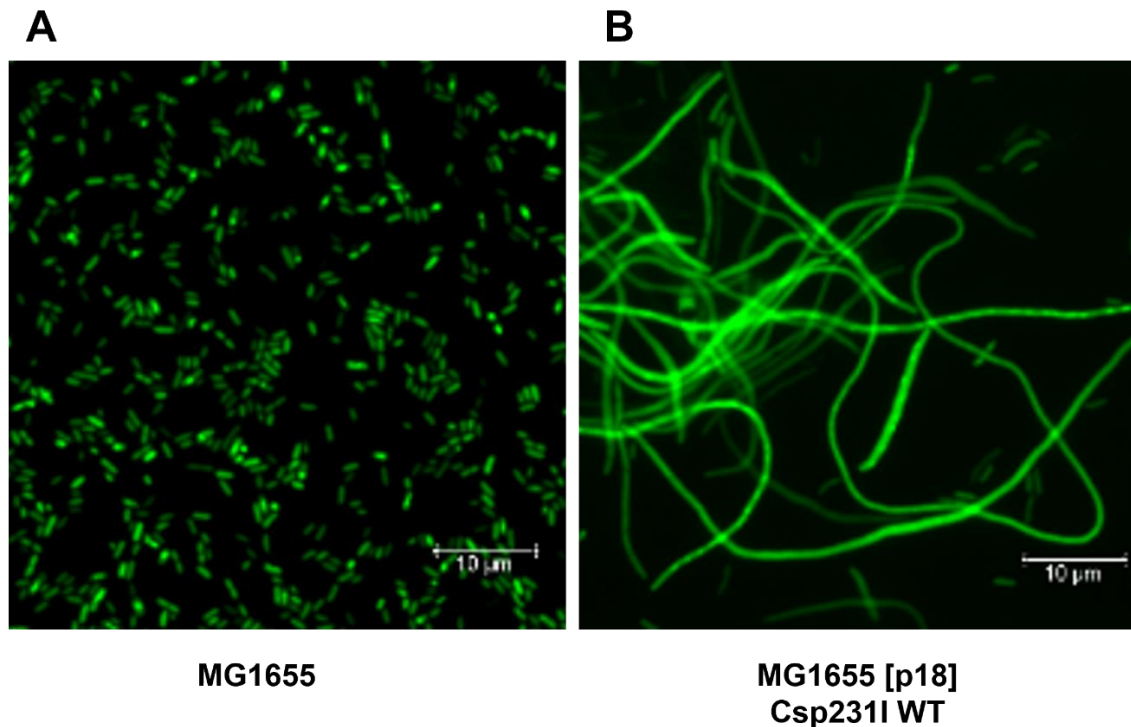
pX = badany konstrukt

4.24. Analizy bioinformatyczne

Nazwa	Charakterystyka	Źródło
WebLogo	Wygenerowanie LOGO, które jest graficznym zestawieniem dopasowania wielu sekwencji aminokwasów lub kwasów nukleinowych	https://weblogo.berkeley.edu/logo.cgi
Protein Molecular Weight	Obliczanie masy cząsteczkowej białka	https://www.bioinformatics.org/sms/prot_mw.html
Multalin toulouse	Porównanie sekwencji kwasów nukleinowych	http://multalin.toulouse.inra.fr/multalin/
ClustalW2	dopasowanie wielu sekwencji aminokwasowych lub kwasów nukleinowych	https://www.ebi.ac.uk/Tools/msa/clustalo/
Bioedit	Porównanie sekwencji aminokwasowych	https://bioedit.software.informer.com/7.2/
Blastn	Wyszukanie podobnych rejonów pomiędzy różnymi sekwencjami nukleotydowymi obliczając istotność statystyczną	https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome
Blastp	Wyszukanie podobnych rejonów pomiędzy różnymi sekwencjami białkowymi obliczając istotność statystyczną	https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=&LINK_LOC=blasttab&LAST_PAGE=blastp
AlphaFold Protein Structure Database	Program do przewidywania struktury białek	https://alphafold.ebi.ac.uk

5. Wyniki

W toku licznych badań w naszym laboratorium nad biologią systemu restrykcyjno-modyfikacyjnego (R-M) Csp231I pochodzącego z bakterii *Citrobacter* sp. RFL231 zaobserwowano ciekawe zjawisko tworzenia się drastycznie wydłużonych komórek *E. coli* pod wpływem wprowadzenia genów tego systemu na plazmidzie (Ryc. 8.). Filamenty te tworzyły się bez względu na warunki hodowli i stosowaną pożywkę.

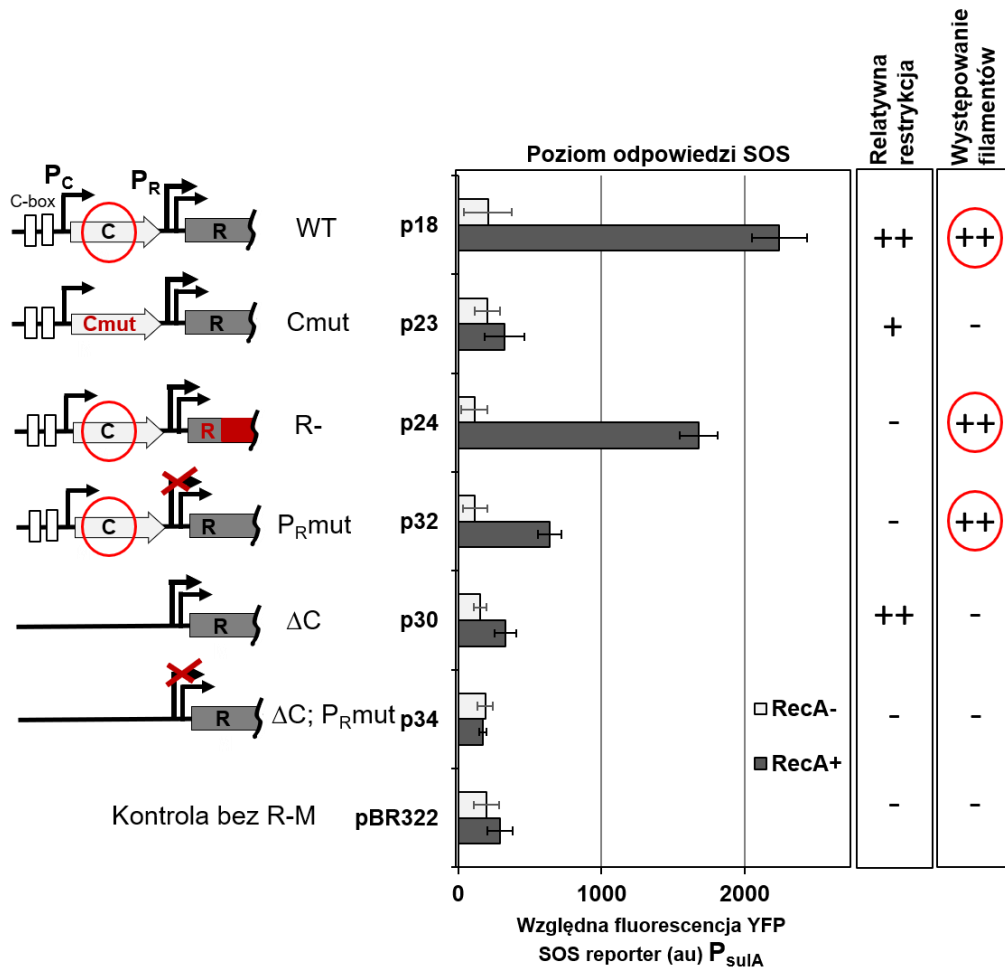


Ryc. 8. Mikroskopia fluorescencyjna komórek *E. coli* MG1655 (**A**) oraz po wprowadzeniu do nich plazmidu p18 niosącego geny systemu R-M Csp231I WT (**B**). Komórki wybarwiono barwnikiem fluorescencyjnym Syto9 (Metody 4.17.) reagującym z kwasami nukleinowymi (Materiały 3.1.).

5.1. Wprowadzenie systemu R-M Csp231I do *E. coli* powoduje filamentację

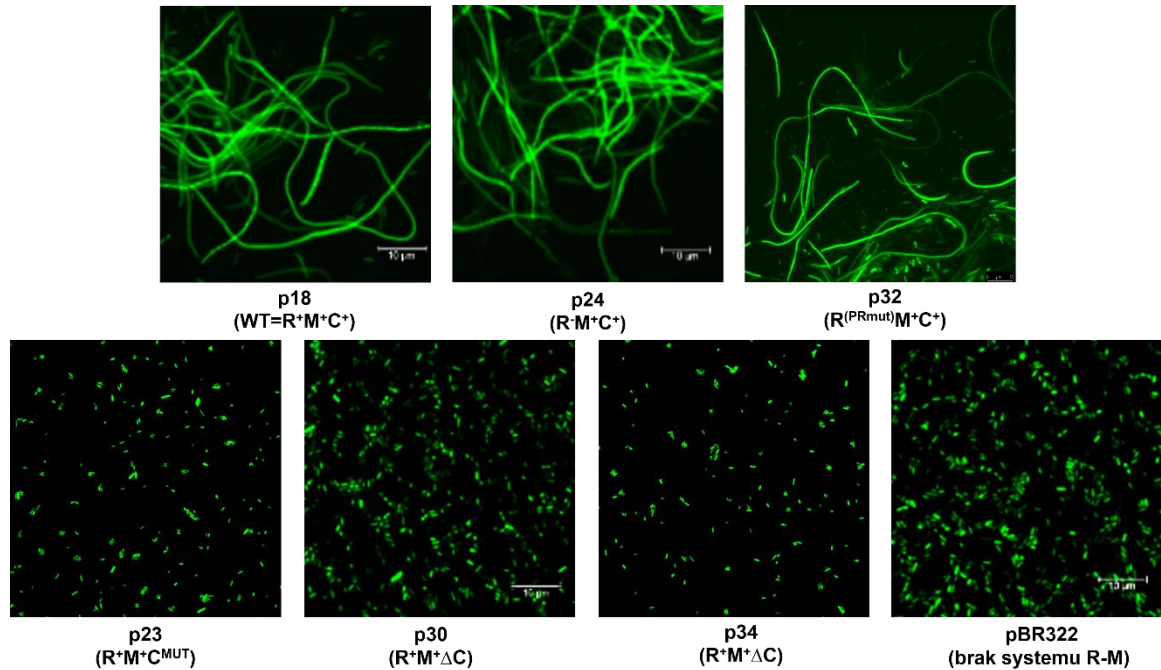
W celu zbadania, który z genów lub elementów genetycznych systemu R-M Csp231I bierze udział w powstawaniu filamentów w komórkach *E. coli*, testowano szereg wariantów tego systemu niesionych przez plazmidy. Warianty te schematycznie przedstawiłam na Rycinie 9 oraz opisałam w Materiały 3.2, w skrócie: plazmid p18 niesie wszystkie naturalne elementy systemu R-M typu WT, p23 niesie zmutowany gen regulatora C, który jest niezdolny do wiązania się z jego naturalną sekwencją specyficzną (tzw. C-box) w promotorze genu C, p24 niesie delecję w genie REazy ($R^{-}M^{+}$), p32 niesie wariant systemu R-M z mutacją w głównym promotorze genu REazy, co skutkuje

obniżoną ekspresją REazy, p30 posiada wariant systemu R-M z całkowitą delecją genu C, co skutkuje podwyższoną ekspresją genu REazy oraz p34, który oprócz delecji genu C (jak p30) posiada też inaktywowany promotor REazy (jak p32). Jako kontrolę użyłam wektor pBR322 bez systemu R-M. Pewne systemy R-M są odpowiedzialne za uruchomienie komórkowej odpowiedzi SOS, z uwagi na produkcję endonukleazy restrykcyjnej, której aktywność może doprowadzać do tzw. auto-restrykcji genomowego DNA. Uszkodzony DNA wywołuje proces RecA-zależnej naprawy, co skutkuje zatrzymaniem podziałów komórkowych i właśnie wydłużaniem komórek bakterii (Handa et al., 2000). Z tego powodu plazmidy z wariantami systemu Csp231I wprowadzono do dwóch typów komórek: *recA*⁺ (szczep MP060) oraz *recA*⁻ (MP064) (Materiały 3.1.; Metody 4.11.). Oba szczepy posiadają w genomie gen reporterowy żółtej fluorescencji pod kontrolą promotora genu *sulA*, aktywnego podczas odpowiedzi SOS.



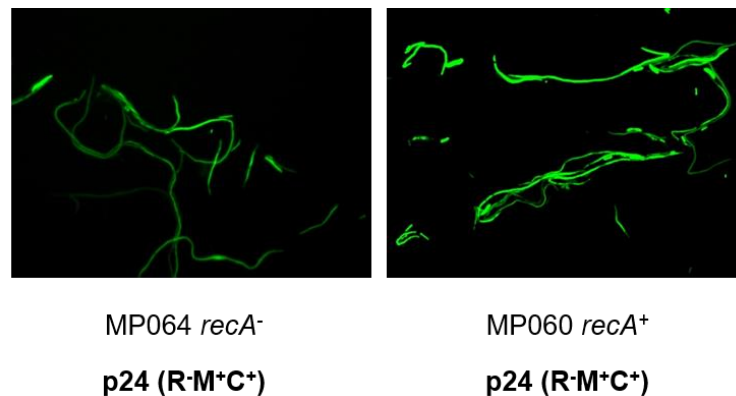
Ryc. 9. Schemat organizacji genetycznej wariantów systemu R-M Csp231I wraz z indukcją odpowiedzi SOS mierzone przez względną fluorescencję w komórkach, które były *recA*⁺ (ciemne słupki) oraz *recA*⁻ (jasne słupki), gdzie gen żółtej fluorescencji jest pod kontrolą promotora *sulA* (P_{sulA-yfp}). Powstawanie filamentów komórkowych (czerwone kółka) jest skorelowane z obecnością aktywnego białka C pochodzącego z systemu R-M Csp231I, a nie z aktywnością restrykcyjną.

Wyniki eksperymentu pokazują, że najwyższy poziom komórkowej odpowiedzi SOS jest skorelowany, nie z najwyższą relatywną restrykcją (Metody 4.23.), a z obecnością aktywnego genu kontrolera C. Tylko w szczepach z aktywnym genem C (system R-M typu dzikiego p18, p24 oraz p32) występują filamenty komórkowe (Ryc. 9 i 10).



Ryc. 10. Mikroskopia fluorescencyjna komórek *E. coli* MG1655WT (Materiały 3.1.) niosących plazmid z genami systemu R-M Csp231I oraz jego wariantami. Górny panel z widocznymi filamentami: p18, p24, p32 oraz dolny bez filamentów: p23, p30, p34, plazmid kontrolny pBR322.

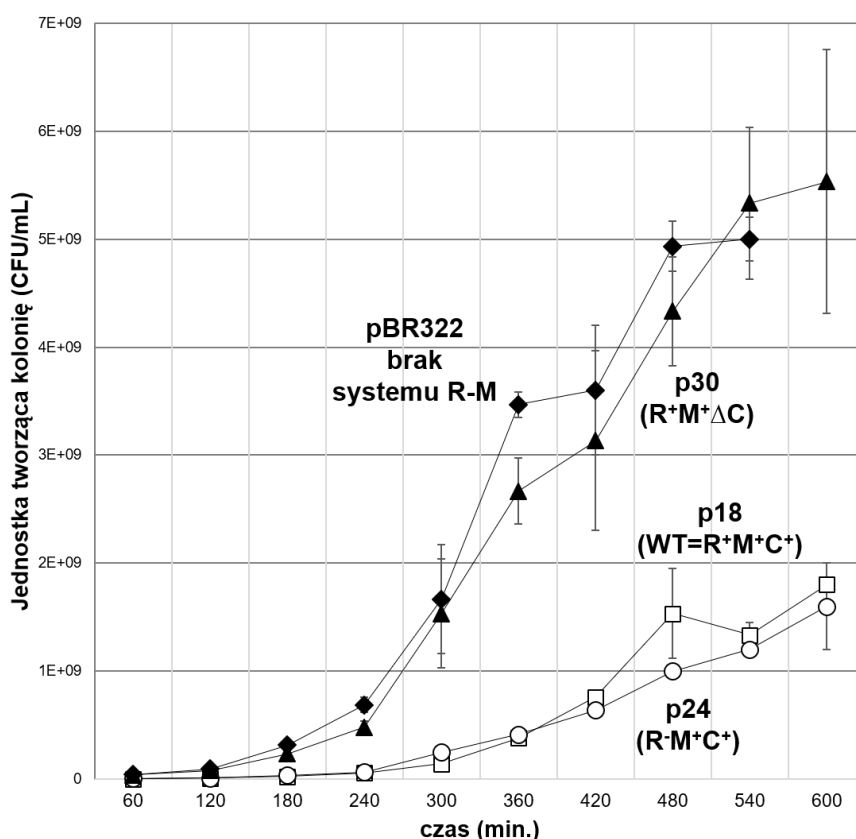
Co ciekawe, filamenty komórkowe występują także w tych samych wariantach w szczepie MP064 *recA*⁻ (Materiały 3.1.). Obserwacje te pozwalają wnioskować, że formacja filamentów nie jest zależna od kompleksu RecA, zatem być może związana jest z innym zjawiskiem niż komórkowa odpowiedź SOS (Ryc. 11.).



Ryc. 11. Mikroskopia fluorescencyjna komórek bakteryjnych MP064 *recA*⁺ (prawy panel) oraz MP064 *recA*⁻ (lewy panel) niosących plazmid z genami systemu R-M Csp231I: p24.

5.2. Wpływ wariantów systemu R-M Csp231I na żywotność komórek *E. coli* MG1655

Kolejno postanowiłam zbadać wpływ poszczególnych wariantów systemu R-M Csp231I, niesionych na plazmidach, na żywotność komórek *E. coli* MG1655WT. Do pomiaru aktywności użyłam konstruktów: p18, który posiada wszystkie naturalne elementy systemu R-M oraz jego wybrane warianty: p24, który ma nieaktywną endonukleazę restrykcyjną oraz p30, który posiada delecję genu C. Jako kontrolę użyłam wariant niosący sam wektor pBR322: bez systemu R-M. Krzywe wzrostu (Ryc. 12.) przygotowałam przy użyciu Metody 4.20.



Ryc. 12. Krzywe wzrostu dla komórek hodowanych w pożywce LB niosących plazmidy: p18 (białe kwadraty), p24 (białe kółka), p30 (czarne trójkąty), pBR322 (czarne romby). Słupki błędów przedstawiają standardowe odchylenie od trzech powtórzeń każdego z nich.

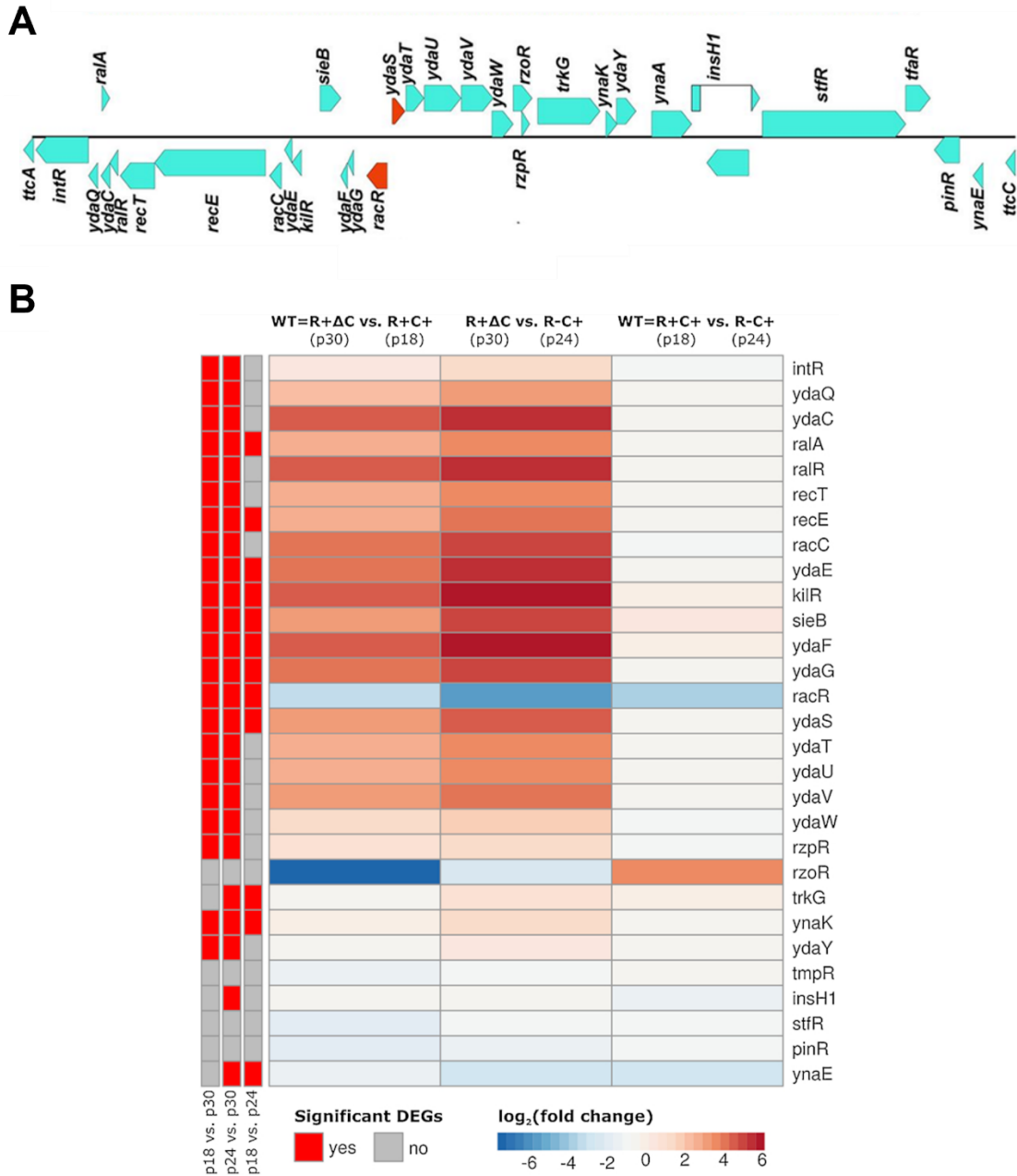
Analizując wyniki zauważyłam, że niższe tempo wzrostu występuje u bakterii niosących plazmidy z aktywnym genem C, co wydaje się być powiązane z występowaniem filamentów komórkowych (Ryc. 9,10.). Niemniej jednak hodowle dla wszystkich testowanych szczepów osiągały fazę stacjonarną i porównywalne OD w tej fazie.

5.3. Analiza transkryptomyczna komórek niosących plazmidy z systemem R-M typu dzikiego oraz jego wariantów

Ustaliłam, że filamentacja jest ściśle zależna od produkcji aktywnego białka regulatorowego C, nawet gdy system R-M jest defektywny pod względem restrykcji DNA (nieaktywna endonukleaza restrykcyjna). Aby zbadać, jakie geny są zaangażowane w ten proces, postanowiliśmy wykonać sekwencjonowanie całkowitego RNA (ang. *RNA-seq*) dla trzech typów komórek, niosących plazmidy: 1) z systemem R-M typu dzikiego p18; 2) z wariantem pozbawionym genu C, ale przy zachowaniu wysokiej relatywnej restrykcji, p30; 3) z wariantem z aktywnym genem C, ale pozbawionym restrykcji DNA, p24.

Celem była analiza transkryptomyczna typu DEG (ang. *differential gene expression*), porównująca poziom produkcji wszystkich transkryptów powstających w komórkach bakterii z wykorzystaniem sekwencjonowania następnej generacji (ang. *Next Generation Sequencing*, NGS). Wszystkie porównawcze analizy transkryptomiczne przeprowadził pan dr Marcin Jąkałski (Negri et al., 2019).

Analiza transkryptomyczna typu DEG (ang. *differential gene expression*) wykazała, że tylko w komórkach z aktywnym genem C poziom ekspresji kilkunastu genów w obrębie rejonu krytycznego faga Rac jest najbardziej podwyższony, nawet 70 razy wobec szczepu z plazmidem z delecją genu C. Pośród kilkunastu genów rejonu Rac o podwyższonej ekspresji, ekspresja tylko jednego genu się obniżyła, i dlatego ten gen, *racR* zainteresował mnie w dalszych analizach (Ryc. 13.).



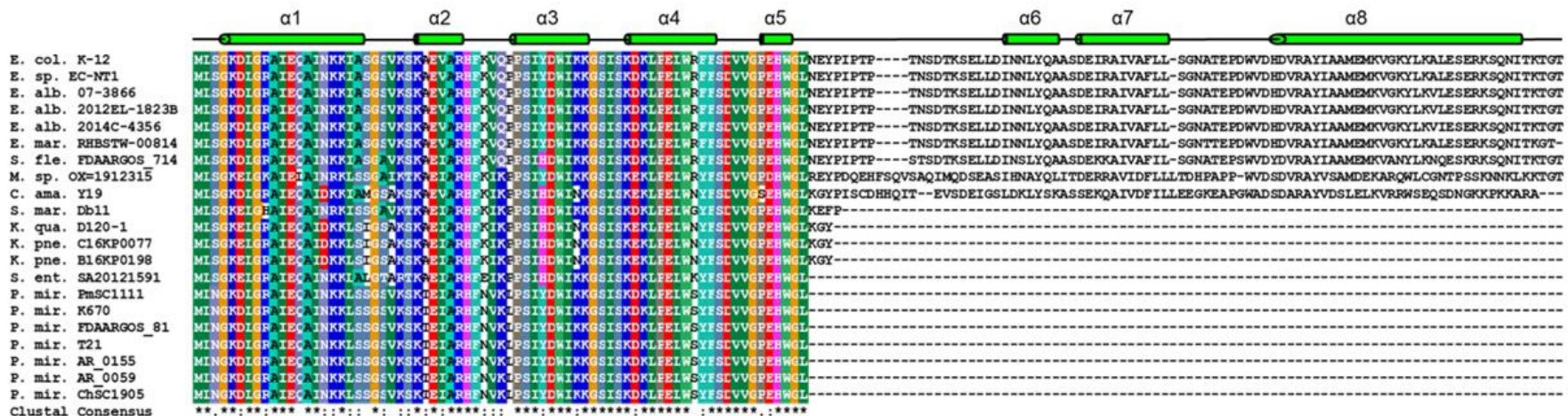
Ryc. 13. A – Mapa genetyczna rejonu profaga Rac położonego w obrębie genu *ttcA* w genomie *E. coli* MG1655. Przepuszczalne geny regulatorowe zaznaczono na pomarańczowo. **B** – Heatmapa wykonana na podstawie analizy typu DEG genów profaga Rac. Geny z podwyższoną ekspresją są pokazane na czerwono, geny z obniżoną regulacją są pokazane na niebiesko. Rycinę opracowałam na podstawie Krishnamurthi et al., 2017 (panel A) oraz Negri et al., 2019 (panel B).

5.4. Analiza *in silico* rejonu faga Rac

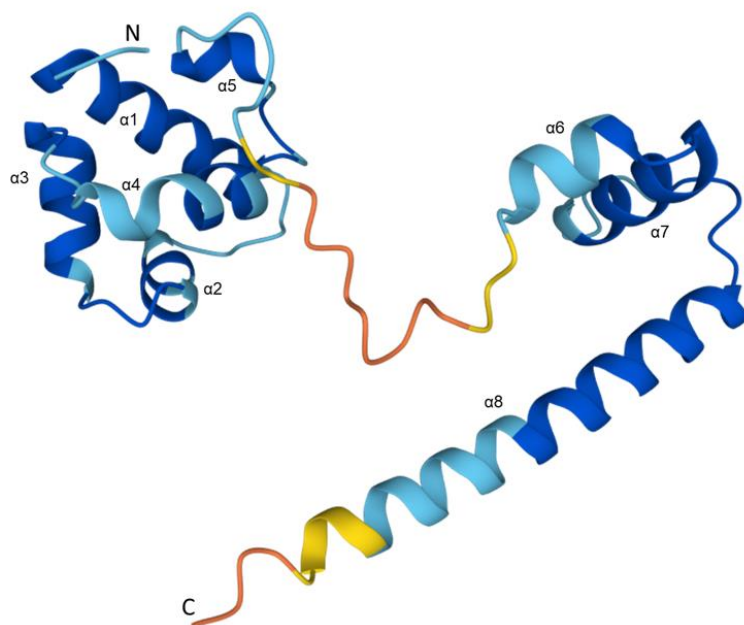
Gen *racR* wchodzi w skład 29 genów obejmujących geny profaga kryptycznego Rac w genomie *E. coli* (Wang et al., 2010) (Ryc. 13 A.). Profagi kryptyczne powstają zwykle podczas przypadkowej mutacji w obrębie genów faga w stanie lizogenii (np. wbudowanego w chromosom bakteryjny). Zazwyczaj profag taki może wycinać się z genomu, ale nie jest zdolny do wytworzenia aktywnych cząstek fagowych. Nie doprowadza również do lizy komórki gospodarza. Wśród tych 29 genów faga Rac występują takie, które nie zostały jeszcze scharakteryzowane. Obszar ten obejmuje około 23 tysięcy par zasad DNA (Wang & Wood, 2016). Analizy komputerowe pozwalają przypuszczać, że białko RacR jest czynnikiem transkrypcyjnym z typowym motywem HTH do interakcji z DNA. Ustalono też, że gen *RacR* należy do grupy genów określanych jako niezbędne do przeżycia (ang. *essential gene*) (Bindal et al., 2017) dlatego, że nie da się otrzymać żywej komórki bakteryjnej pozbawionej tego genu (Baba et al., 2006). Największa delecja niepowodująca śmierci komórki jest możliwa z zachowaniem genu *racR* oraz *ydaS* wraz z regionem intergenowym występującym pomiędzy tymi genami (Krishnamurthi et al., 2017).

Przy użyciu programu Blast, znalazłam homologi białka RacR, licznie występujące w rodzinie *Enterobacteriaceae* i wykonałam analizę porównawczą dla sekwencji białkowej w programie MultAlin (Metody 4.24.) (Ryc. 14.). Analiza ta wykazała występowanie białka RacR w dwóch formach: dłuższej (8 helis) i krótszej (5 helis). Dłuższy wariant występuje w rodzajach: *Escherichia*, *Morganella* i *Shigella*, a krótszy w rodzajach: *Proteus*, *Salmonella*, *Serratia* i *Klebsiella*.

Kolejnym krokiem było wygenerowanie struktury monomeru regulatora RacR z *E. coli* przy użyciu programu AlphaFold (Metody 4.24.). Monomer ten zbudowany jest z 8 helis. Białko to posiada nieustrukturyzowany łącznik zbudowany z 18 reszt aminokwasowych, który łączy dwie domeny białkowe. N-końcowa domena zawiera 5 helis, które są obecne we wszystkich badanych przeze mnie homologach RacR i C-końcowa domena zbudowana z 3 helis obecna tylko u niektórych gatunków (Ryc. 15.). Wydaje się, że krótsza forma RacR jest ograniczona do obecności tylko jednej domeny bez łącznika.



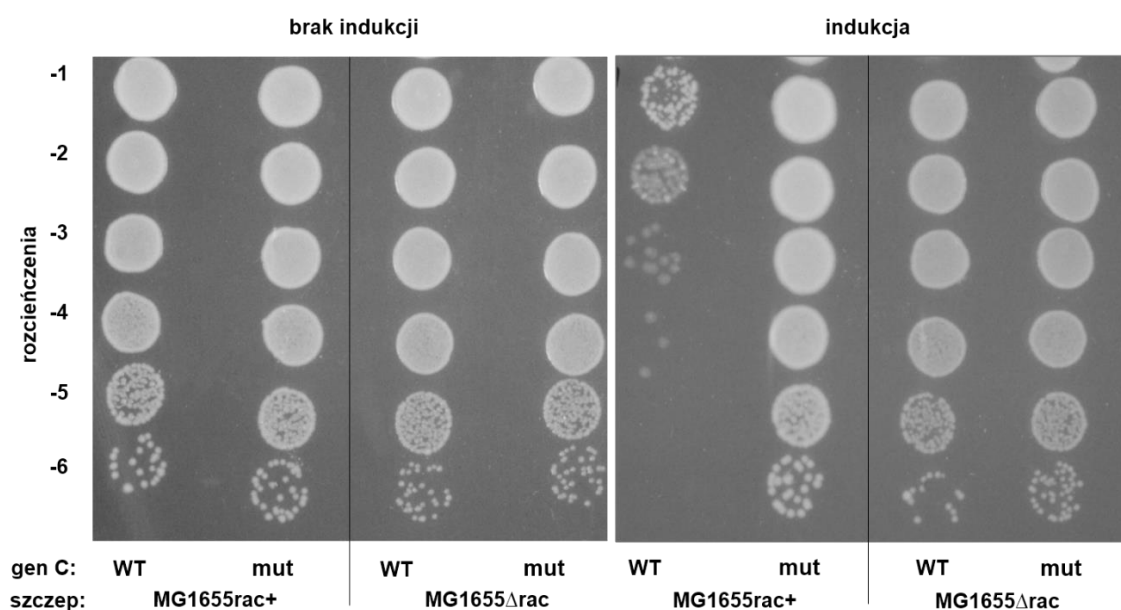
Ryc. 14. Porównanie struktury I-rzędowej białka RacR z *Escherichia coli* K-12 (MG1655) i jego homologów w rodzinie *Enterobacteriaceae*. Wykorzystane skróty: E.col. (*Escherichia coli*), E.alb. (*Escherichia albertii*), E.sp. (*Enterobacter* sp.), E.mar. (*Escherichia marmotae*), M.sp. (*Morganella* sp.), C.ama. (*Citrobacter amalonaticus*), S.flex. (*Shigella flexneri*), P.mir. (*Proteus mirabilis*), K.qua. (*Klebsiella quasipneumonia*), K.pne. (*Klebsiella pneumonia*), S.mar. (*Serratia marcescens*), S.ent. (*Salmonella enterica*).



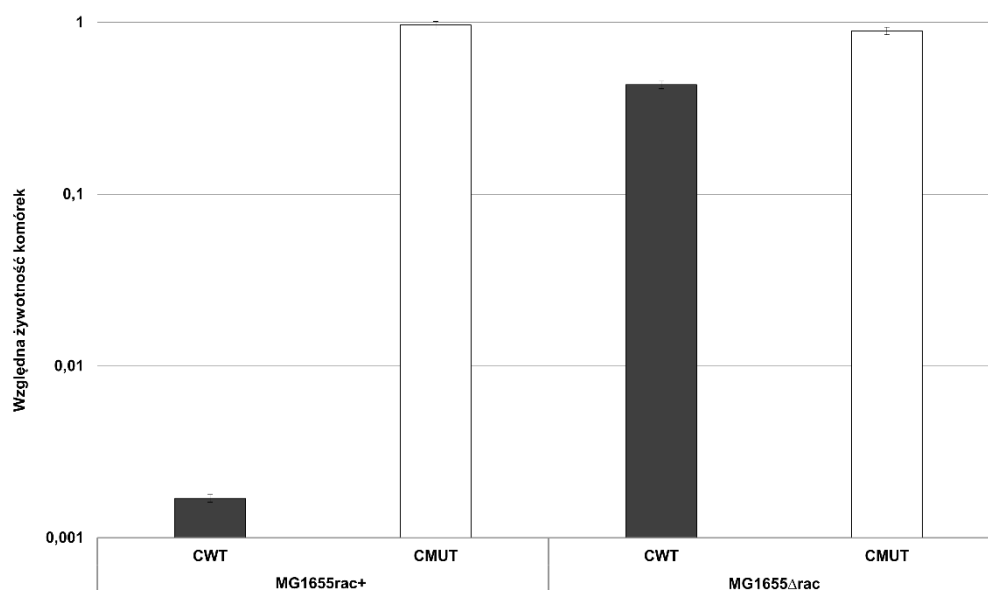
Ryc. 15. Struktura drugorzędowa monomeru białka RacR. Białko RacR złożone jest z 158 reszt aminokwasowych. Posiada 8 alfa-helis oznaczonych na rysunku. Model został wygenerowany za pomocą programu AlphaFold (Metody 4.24.). Kolor ciemnoniebieski wskazuje na największe prawdopodobieństwo przypisania danego motywu białkowego, najmniejsze – kolor żółty i pomarańczowy.

5.5. Badanie wpływu genu C na żywotność komórek bakteryjnych w szczepie posiadającym rejon profaga Rac w genomie

Następnie postanowiłam sprawdzić, jak indukowana ekspresja genu C wpływa na żywotność komórek bakteryjnych, które posiadają w swoim genomie profaga Rac, czyli w szczepie MG1655WT (Materiały 3.1.). Jako kontrolę użyłam szczep z delecją tego rejonu (MG1655 Δ rac). Do obu wyżej wymienionych szczepów wprowadziłam plazmid, gdzie ekspresja genu C jest pod kontrolą promotora P_{BAD} : pBAD-CWT (typ dziki genu C) lub pBAD-Cmut, (Materiały 3.2.), gdzie gen Cmut produkuje wariant białka C niezdolnego do wiązania się ze swoimi operatorami, tzw. sekwencją C-box. (Materiały 3.2.). Sekwencja C-box w formie naturalnej jest nieobecna w genomie i nieobecna na plazmidzie pBAD-CWT. Promotor P_{BAD} umożliwia indukcję ekspresji genu C za pomocą 0,1% arabinozy, a hamowanie ekspresji odbywa się pod wpływem glukozy (Guzman et al., 1995). Hodowle bakteryjne w LB w fazie logarytmicznej indukowałam arabinozą (Metody 4.18.), a następnie komórki rozcieńczałam seryjnie i nakrapiałam na podłoże LB-agar z arabinozą. Otrzymane wyniki przedstawiłam na Rycinie 16.



Ryc. 16. Żywotność komórek bakteryjnych w obecności białka C w komórkach: MG1655WT(*rac*+) i MG1655Δ*rac*. Wykorzystałam plazmid niosący promotor arabinozowy do indukcji ekspresji genu C typu dzikiego: pBAD-CWT (Materiały 3.2.) lub jego wariantu: pBAD-Cmut (Materiały 3.2.).

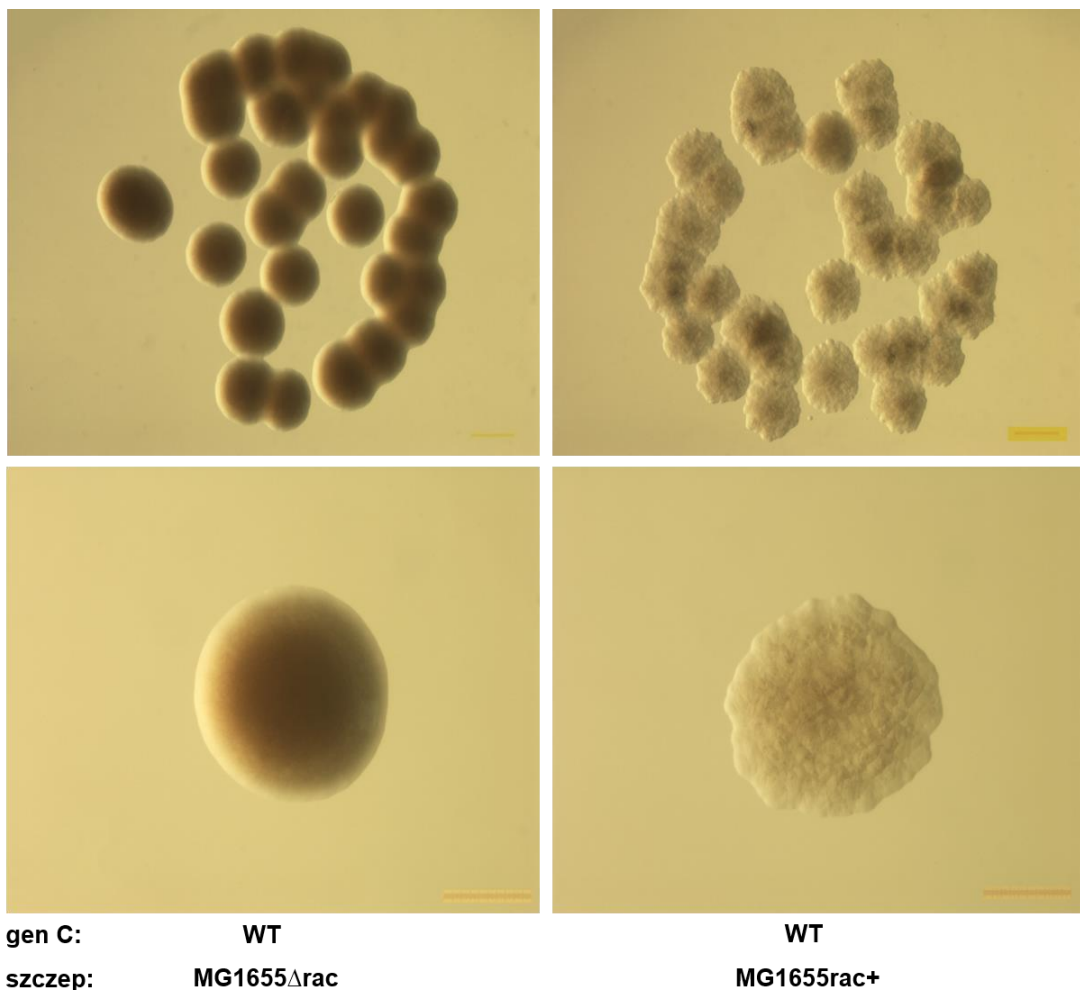


Ryc. 17. Żywotność komórek bakteryjnych *E. coli* w dwóch szczepach: MG1655WT(*rac*+) i MG1655Δ*rac*. Szare słupki pokazują wpływ białka typu dzikiego: CWT, a białe - Cmut. Wyniki opracowano w formie ilościowej na podstawie ryciny 16 (w powtórzeniach). Względna żywotność komórek została określona w taki sposób, że CFU dla MG1655WT CMUT uznałam za 100%.

Wyniki tych doświadczeń: podejście jakościowe (Ryc. 16.) i ilościowe (Ryc. 17.) wskazują jednoznacznie, że komórki bakteryjne prezentują znacznie obniżoną przeżywalność tylko w szczepie *rac*+ z aktywnym genem C (ok. 1000-krotnie mniejszą),

co sugeruje możliwą interakcję białka C z nieznaną sekwencją DNA w rejonie faga Rac. Ta interakcja skutkuje negatywnie na żywotność komórki gospodarza. Efektu tego nie obserwuje się w komórkach Δ rac.

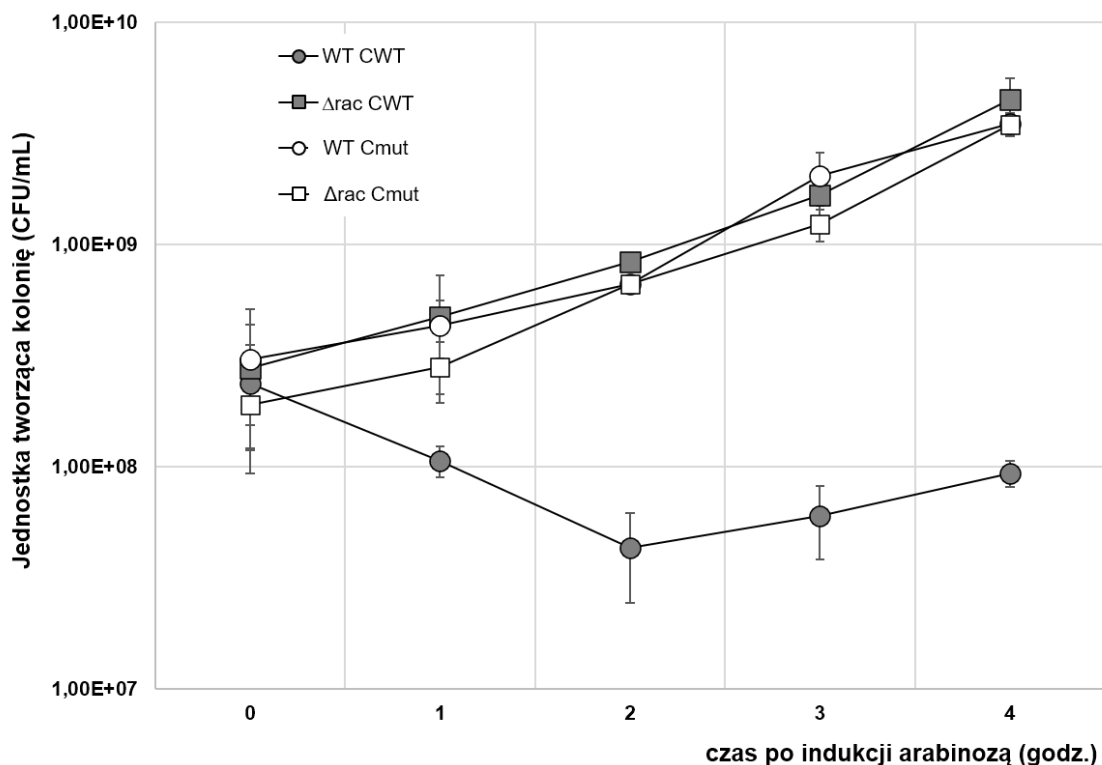
Zauważyłam również nietypową morfologię kolonii bakteryjnych o obniżonej żywotności: w szczepie MG1655WT z plazmidem niosącym gen białka C typu dzikiego: pBAD-CWT. Objawiał się on poszarpaną linią brzegową kolonii, nieregularną strukturą całej kolonii bakteryjnej, mniejszą gęstością i przez to większą przejrzystością (Ryc. 18.). W badaniu mikroskopowym potwierdzono drastyczną filamentację komórek w takiej zmienionej kolonii.



Ryc. 18. Obrazy mikroskopowe ukazujące fenotyp otrzymanych kolonii bakteryjnych. Nietypowa morfologia pojedynczej kolonii bakteryjnej składająca się z komórek filamentujących (aktywny gen C / rac+: prawy panel) w porównaniu do normalnej kolonii kontrolnej (aktywny gen C / rac-: lewy panel).

Następnie postanowiłam sprawdzić żywotność komórek w czasie w hodowli płynnej dla tych samych dwóch szczepów bakteryjnych: MG1655WT oraz MG1655 Δ rac z dwoma wariantami białka C: pBAD-CWT oraz pBAD-Cmut (Materiały 3.2.). Hodowlę

doprowadzono do wczesnej fazy logarytmicznej, indukowano ekspresję genu C arabinozą i pobierano próbki czasowe, co 1 godzinę. Jak oczekiwano, szczep z niską żywotnością (rac^+ / CWT) odznaczał się spadkiem CFU już po godzinie od indukcji ekspresji C (Ryc. 19.). Co ciekawe, obserwacja w warunkach dłuższej indukcji wykazała, że część komórek odzyskała swój normalny fenotyp (brak filamentacji w obrazowaniu mikroskopowym) i nawet zaobserwowałam niewielki wzrost żywotności po ok. 4 godz. od indukcji. Wyniki te wskazują, że efekt filamentacji oraz osłabionej żywotności może być zjawiskiem odwracalnym, a być może wynika z wyczerpania zasobów arabinozy w pożywce.

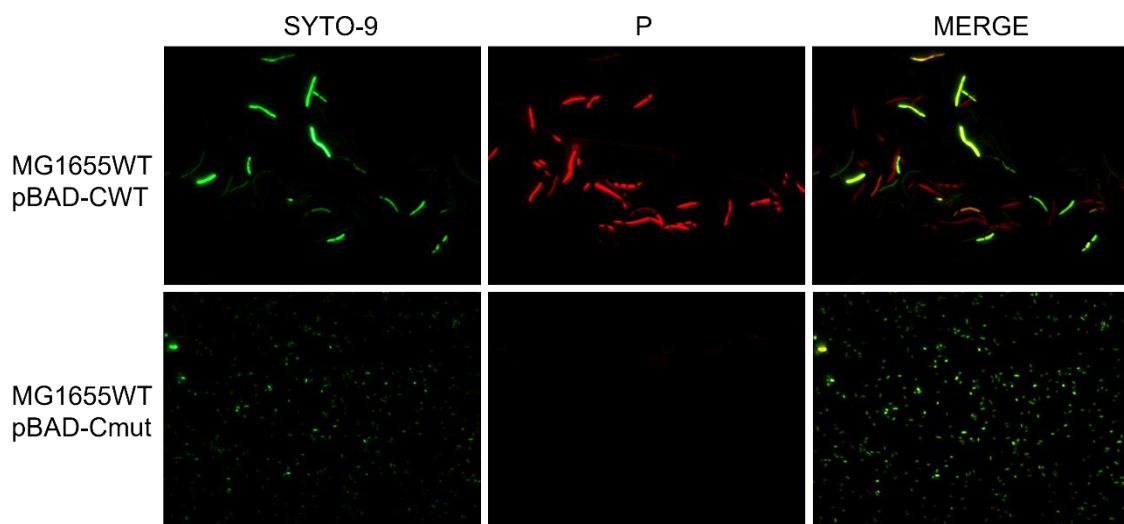


Ryc. 19. Żywotność komórek bakteryjnych *E. coli* w dwóch szczepach: MG1655WT (rac^+) i MG1655 Δrac podczas indukcji białka C. Białka CWT lub Cmut indukowałam 0,01 % arabinozą. Zmierzyłam jednostkę tworzącą kolonię (CFU) dla czterech kombinacji hodowli bakteryjnych: MG1655WT/CWT; MG1655 Δrac /CWT; MG1655WT/Cmut; MG1655 Δrac /Cmut. Pomiary zbierałam przed indukcją (czas 0) i co godzinę po indukcji. Słupki błędów pokazują standardowe odchylenie mierzone dla trzech równoległych hodowli.

5.6. Mikroskopia fluorescencyjna LIVE/DEAD komórek bakteryjnych MG1655WT niosących białko C na plazmidzie

Postanowiłam zastosować barwienie fluorescencyjne LIVE/DEAD (Metody 4.16.) komórek bakteryjnych MG1655WT zawierających plazmid pBAD-CWT oraz

pBAD-Cmut – kontrolnie, aby sprawdzić, czy powstające filamenty komórkowe są żywe, czy martwe. Żywotność badałam w obecności wysokiego stężenia białka C (po indukcji arabinozą). Analizę przeprowadziłam za pomocą barwienia fluorescencyjnego LIVE/DEAD BacLight (Thermo Fisher) (Metody 4.12.). Zestaw ten zawiera dwa barwniki: SYTO®9 oraz jodek propidyny – PI (ang. *propidium iodide*). SYTO®9 o małej masie cząsteczkowej przenika do wnętrza komórki przez spójne błony cytoplazmatyczne, a PI o dużej masie cząsteczkowej przenika tylko do tych z uszkodzeniami w błonach. Zastosowanie zestawu LIVE/DEAD® umożliwia zróżnicowanie populacji bakterii na komórki żywe i martwe pod względem różnic w ciągłości błon cytoplazmatycznych. Badanie to pozwoliło mi zauważyć, że powstałe filamenty komórkowe mają charakter mieszany – są zarówno żywe, jak i martwe. Ponadto w niektórych przypadkach widać, że nawet pojedynczy filament nie jest tylko martwy lub żywy, ale też ma charakter mieszany w obrębie segmentów długiego filamentu (Ryc. 20.).

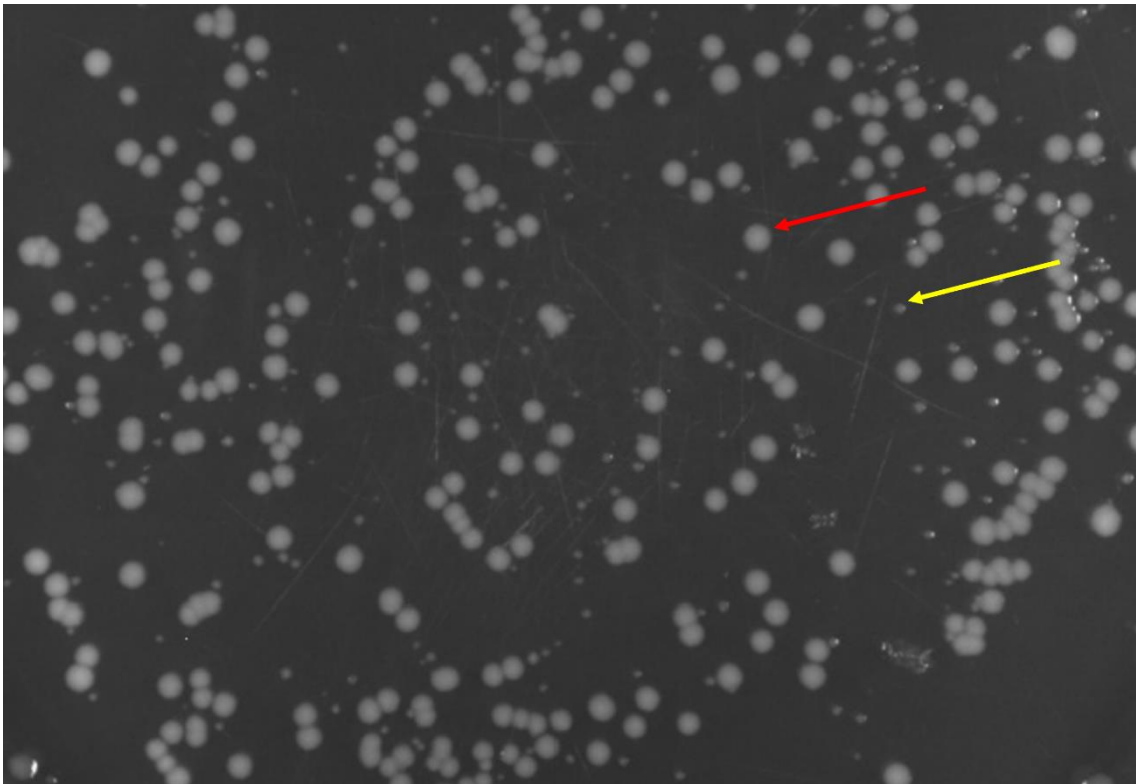


Ryc. 20. Mikroskopia fluorescencyjna LIVE/DEAD bakterii posiadających białko C lub jego zmutowany wariant jako kontrola. Powstałe filamenty bakteryjne przy aktywnym białku C są zarówno żywe, jak i martwe (górny panel) wobec kontroli z wariantem białka Cmut (obecne tylko żywe komórki (dolny panel).

5.7. Badanie fenotypu filamentacyjnego w czasie dla komórek z aktywnym genem regulatora C

Zaobserwowałam, że efekt powstawania filamentów w czasie jest relatywnie odwracalny, zatem ciekawiło mnie, jak ten efekt będzie przebiegał w kontekście wielu generacji bakterii. Żeby uniknąć induktora ekspresji genu C, użyłam plazmidu p24 (C+R-M+), który niesie gen C pod kontrolą naturalnego konstytutywnego promotora. Plazmid

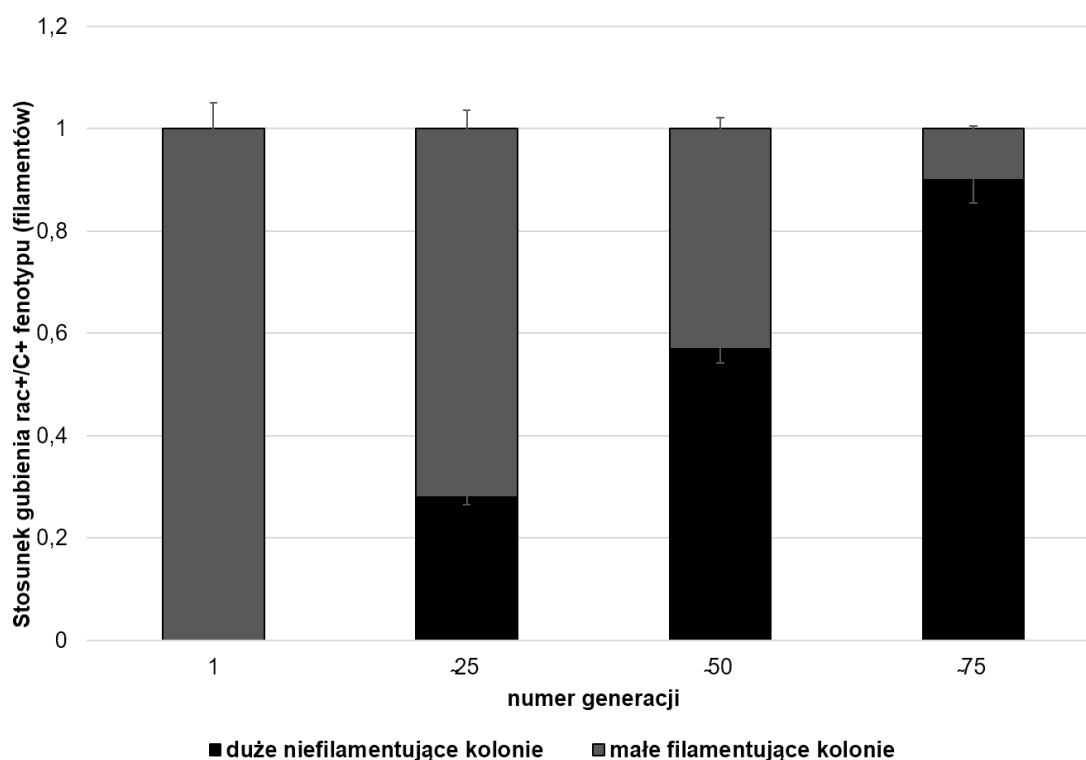
ten wprowadziłam do komórek MG1655WT rac+ (Materiały 3.1.). Po otrzymaniu transformantów komórki pasażowałam do świeżej pożywki, co ok. 25 pokoleń stosując presję selekcyjną w postaci antybiotykoodporności niesionej przez plazmid. Na początku eksperymentu wszystkie komórki były filamentujące i tworzyły małe kolonie. Z czasem zaobserwowałam dwa różne typy morfologii: małe, filamentujące i duże o normalnej morfologii (Ryc. 21.).



Ryc. 21. Obraz przedstawiający pasażowane kolonie wysiewane na płytkę LB-agar z odpowiednim antybiotykiem. Żółtą strzałką oznaczyłam kolonie małą, filamentującą charakterystyczną dla szczepu MG1655WT z aktywnym genem C, natomiast czerwoną strzałką dużą niefilamentującą kolonię.

Następnie przeliczałam stosunek kolonii o tych dwóch morfologiach w procentach w zależności od generacji bakterii. Tendencja pokazywała silny wzrost liczby komórek dużych – niefilamentujących w stosunku do małych – filamentujących. W ok. 75 generacji komórek, większość z nich już utraciła filamentacyjny charakter (Ryc. 22.). Zatem badanie wykazało, że w komórkach jest silna presja, aby pozbyć się przyczyny formowania się filamentów. Jeśli przyczyna ta jest na poziomie genetycznym, być może dochodzi do powstawania spontanicznych mutacji, które neutralizują to zjawisko. Dlatego postanowiliśmy sprawdzić, czy na końcu eksperymentu gen C niesiony przez plazmid jest ciągle aktywny. Plazmid izolowałam z dużych kolonii i ponownie transformowałam szczep MG1655 rac+. Za każdym razem uzyskiwałam transformanty z

komórkami filamentującymi, co wskazuje, że gen C nie uległ mutacji, a potencjalne mutacje supresorowe mogły pojawić się w genomie.

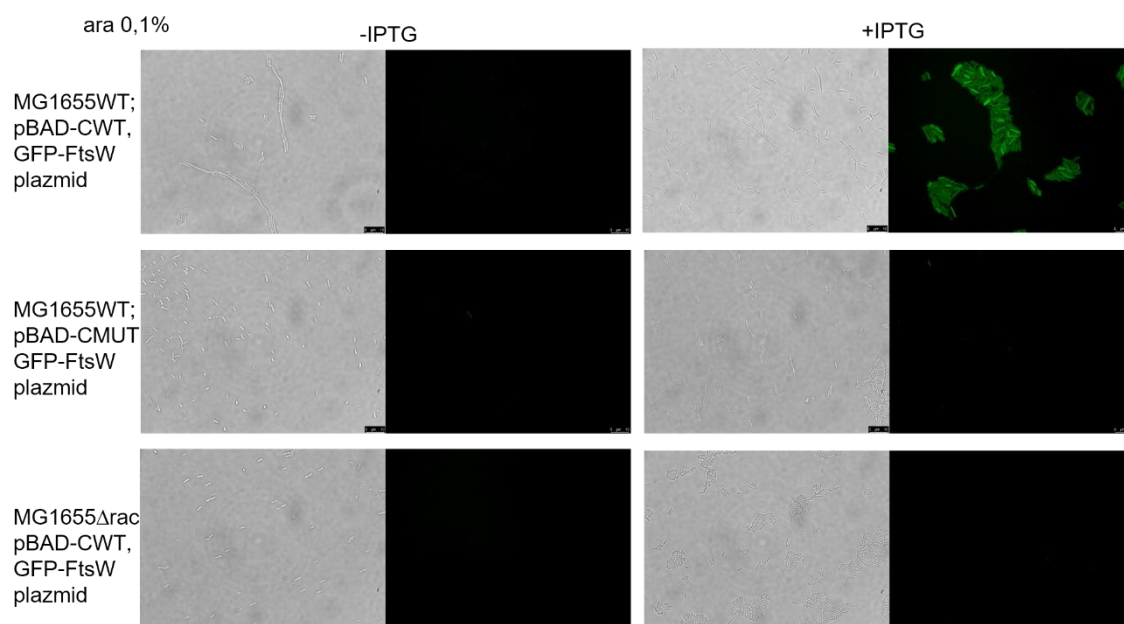


Ryc. 22. Utrata fenotypu filamentacyjnego w kolejnych generacjach: 1, 25, 50 i 70. Kolor szary – małe filamentujące kolonie, czarny – duże niefilamentujące.

5.8. Badanie ekspresji białka FtsW w komórkach filamentujących

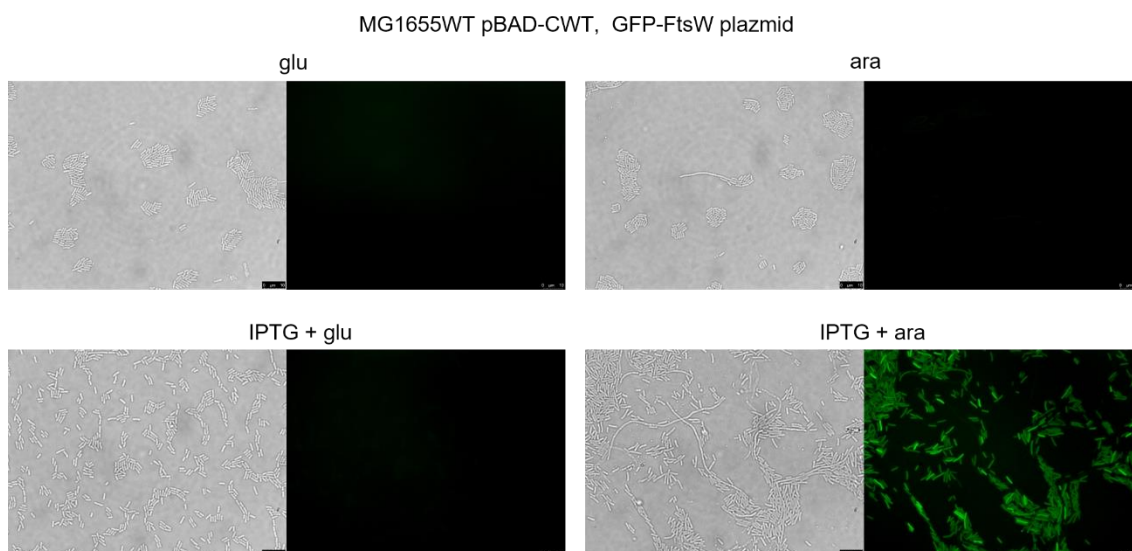
Podział komórki bakteryjnej zachodzi w jej centralnej części i następuje po zduplikowaniu DNA i rozdzieleniu go na dwa nukleoidy. Proces podziału rozpoczyna się od zlokalizowania FtsZ w centrum komórki i utworzenia przegrodowej struktury pierścieniowej zwanej Z-ringiem. Gdy przegroda (septa) jest w pełni uformowana, pierścień Z znika, a komórki potomne oddzielają się. Za tym procesem podąża uporządkowany zespół innych białek (m.in.: FtsA, ZipA, FtsK, FtsQ, FtsL, FtsB, FtsW, FtsI, FtsN) (Hugonnet et al., 2016). Aby sprawdzić możliwą wzmożoną aktywność podziałów komórkowych w badanym przeze mnie modelu postanowiłam wykorzystać białko FtsW biorące udział w podziale komórki na późnym etapie (Craven et al., 2022). Do tego celu wykorzystałam plazmid pDSW360 (Materiały 3.2.) posiadający gen *ftsW* w fuzji z genem GFP, który indukowałam za pomocą IPTG, aby uzyskać ekspresję tego białka fuzyjnego. Dużą aktywność białka fluorescencyjnego GFP zauważyłam jedynie w

wariancie, który znajdował się w dzikim szczepie MG1655 (Materiały 3.1.) zawierającym pBAD-CWT (Materiały 3.2.) oraz plazmid pDSW360 (Materiały 3.2.). Zatem fluorescencja, a co za tym idzie ekspresja białka FtsW jest widoczna tylko przy aktywnym białku C w kombinacji w obecności rejonu *rac* (Ryc. 23.).

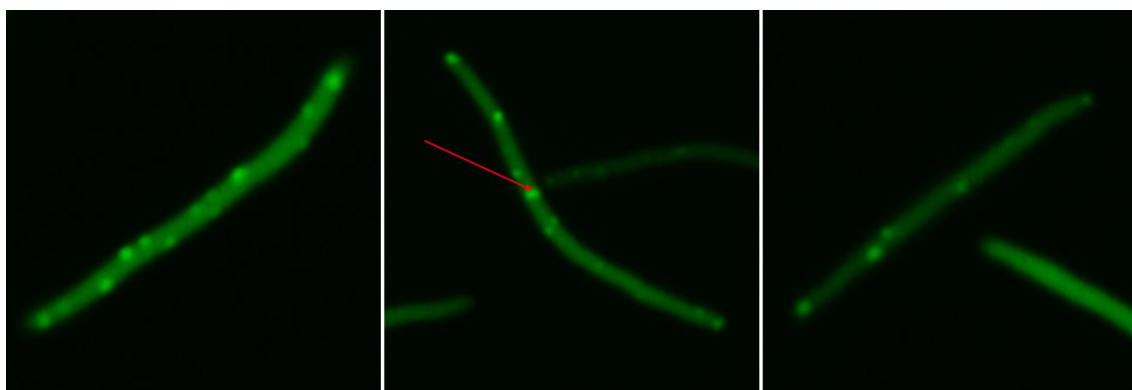


Ryc. 23. Mikroskopia fluorescencyjna komórek bakteryjnych MG1655WT lub MG1655Δrac z wariantami plazmidów pBAD-CWT lub pBAD-Cmut i plazmidem syntetyzującym białko FtsW (pDSW360) indukowane za pomocą IPTG.

W celu upewnienia się, czy efekt widocznej fluorescencji GFP nie jest wynikiem indukcji, któregośkolwiek innego czynnika występującego w tym modelu wykonałam kolejną mikroskopię fluorescencyjną. Przygotowałam jeden wariant, ten który wykazywał zjawisko fluorescencji, czyli MG1655WT z plazmidami pBAD-CWT oraz pDSW360 (Materiały 3.2.) i przygotowałam z niego 4 osobne hodowle, do których dodawałam: glukozę, arabinozę, IPTG + glukozę i IPTG + arabinozę. Następnie przygotowałam preparaty do mikroskopii fluorescencyjnej. Wynik był jednoznaczny, mianowicie fluorescencja białka GFP::FtsW występowała tylko w obecności aktywnego białka C oraz przy obecności rejonu *rac*. Fluorescencji nie zaobserwowałam przy dodaniu samej glukozy, arabinozy, czy nawet IPTG z glukozą (Ryc. 24.). Co ciekawe, przyglądając się tym komórkom w powiększeniu zauważyłam ciekawe zjawisko, mianowicie akumulację białka GFP w pewnych miejscach wydłużonej komórki bakteryjnej (Ryc. 25.), które możliwie wskazują miejsce inicjacji tworzenia się septy, która ostatecznie nie doprowadza do rozdzielenia potomnych komórek. Możliwe, że to zaburzenia ma charakter defektu regulacji tego procesu.

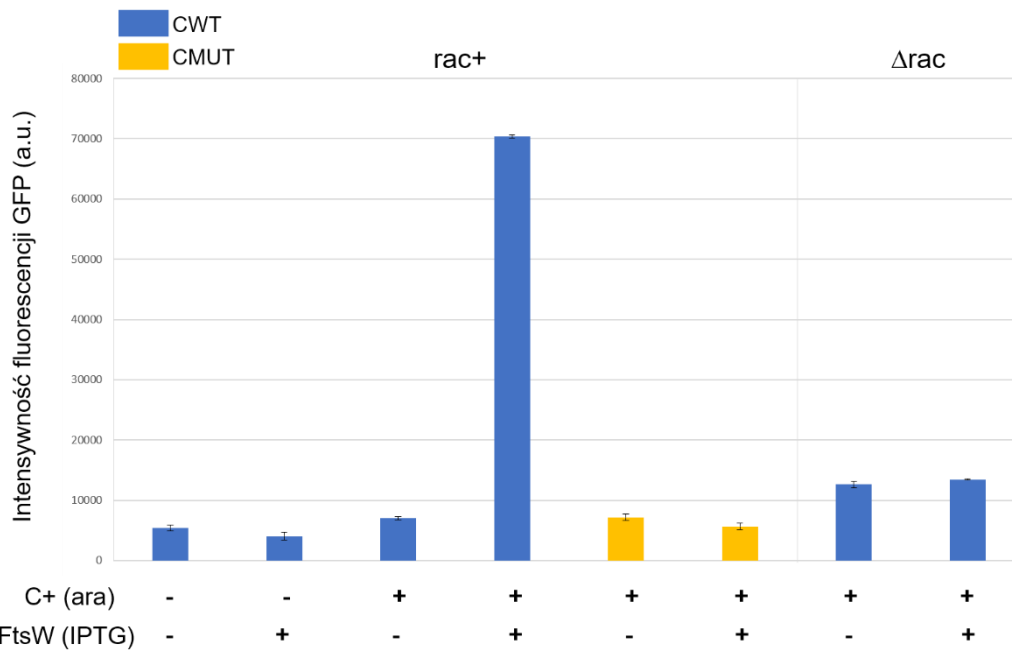


Ryc. 24. Mikroskopia fluorescencyjna komórek bakteryjnych MG1655WT (Materiały 3.1.) z plazmidem pBAD-CWT (Materiały 3.2.) oraz GFP-FtsW (pDSW360) (Materiały 3.2.) po dodaniu induktorów.



Ryc. 25. Mikroskopia fluorescencyjna wydłużonych komórek bakteryjnych MG1655WT (Materiały 3.1.) posiadających rejon rac z aktywnym genem C indukowanym za pomocą arabinozy oraz aktywnym białkiem FtsW indukowanym za pomocą IPTG. Akumulacje GFP zaobserwowałam tylko w komórkach filamentujących.

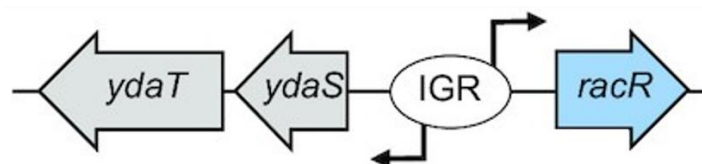
Kolejno postanowiłam zmierzyć intensywność fluorescencji GFP. Aby to zrobić hodowałam komórki MG1655WT i MG1655 Δ rac przez noc z wytrząsaniem, a następnie delikatnie osadzałam i zawieszałam w soli fizjologicznej. Mierzyłam poziom fluorescencji zebranych próbek. Wyniki przedstawiłam w postaci wykresu słupkowego. Tylko aktywny gen C jest w stanie indukować wysoką ekspresję genu białka fuzyjnego GFP::FtsW (Ryc. 26.).



Ryc. 26. Poziom białka fuzyjnego GFP::FtsW wyrażony intensywnością zielonej fluorescencji. Indukcja GFP dla komórek niosących plazmidy (+) lub ich braku (-) mierzylam za pomocą względnej fluorescencji w kontekście: CWT (niebieskie słupki) i Cmut (żółte słupki).

FtsW jest integralnym białkiem błonowym wymaganym do podziału komórek *Escherichia coli* (Boyle et al., 1997) i przypuszczalnie również u innych bakterii. Jego dokładna rola biochemiczna nie jest jeszcze znana. FtsW jest jednym z dziewięciu podstawowych białek podziału komórkowego u *E. coli*. Wszystkie z nich lokalizują się w miejscu podziału (Mercer & Weiss, 2002). Wydaje się, że w badanych przeze mnie filamentach podział komórkowy jest zaburzony. Stąd też możliwa jest akumulacja białka FtsW w miejscu septy, jednak, komórki się nie dzielą, jak pokazałam na rycinie 25. Nie jest do końca jasne, dlaczego wysoki poziom białka FtsW jest zauważalny tylko przy aktywnym genie C.

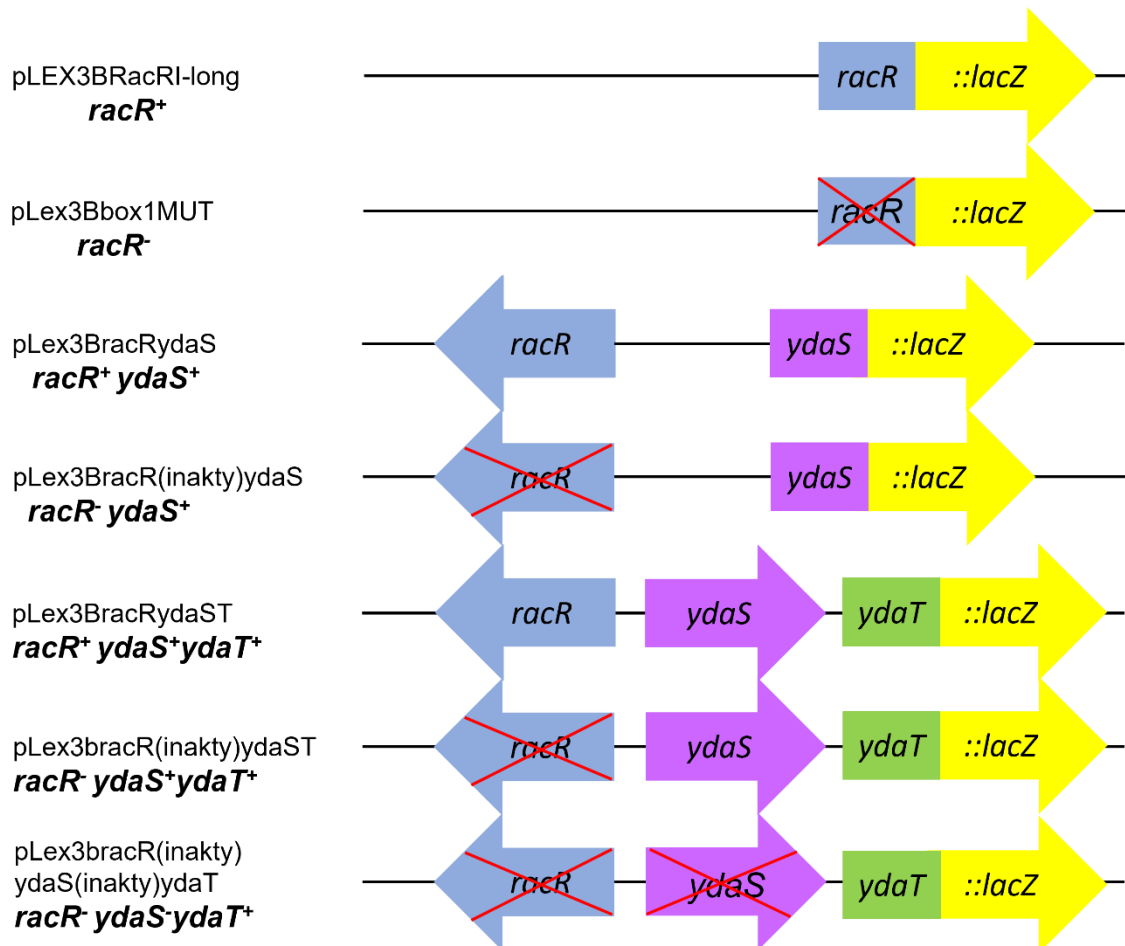
5.9. Wpływ białka C na ekspresję genów *racR*, *ydaS* oraz *ydaT*



Ryc. 27. Schematyczna mapa genetyczna operonu *ydaST* oraz *racR* wraz z regionem intergenowym (IGR).

Postanowiłam sprawdzić wpływ białka C na aktywność transkrypcyjną genów *racR*, *ydaS* i *ydaT* (Ryc. 27.) w warunkach wzrastającej ilości białka C w komórce wykorzystując w tym celu promotor arabinozowy: od braku C (glukoza) do akumulacji C (0,1% arabinoza). Równoległe zaplanowałam badanie wpływu wariantu Cmut (kontrola negatywna).

Do tego doświadczenia wybrałam komórki MG1655Δ*rac*, aby uniknąć wpływu genu *racR* / *ydaS* / *ydaT* znajdujących się w genomie. Do tej pory nie określono położenia ani aktywności promotorów tych genów. Komórki transformowałam najpierw plazmidem pBAD-CWT lub pBAD-Cmut, a potem wprowadzałam plazmid z odpowiednim wariantem genów *racR*, *ydaS* i *ydaT* w fuzji z genem reporterowym *lacZ*, który zapewnia ilościowy pomiar ekspresji genów. W tym celu wykonałam serię konstruktów, którą przedstawiłam na rycinie 28 (Materiały 3.2.).

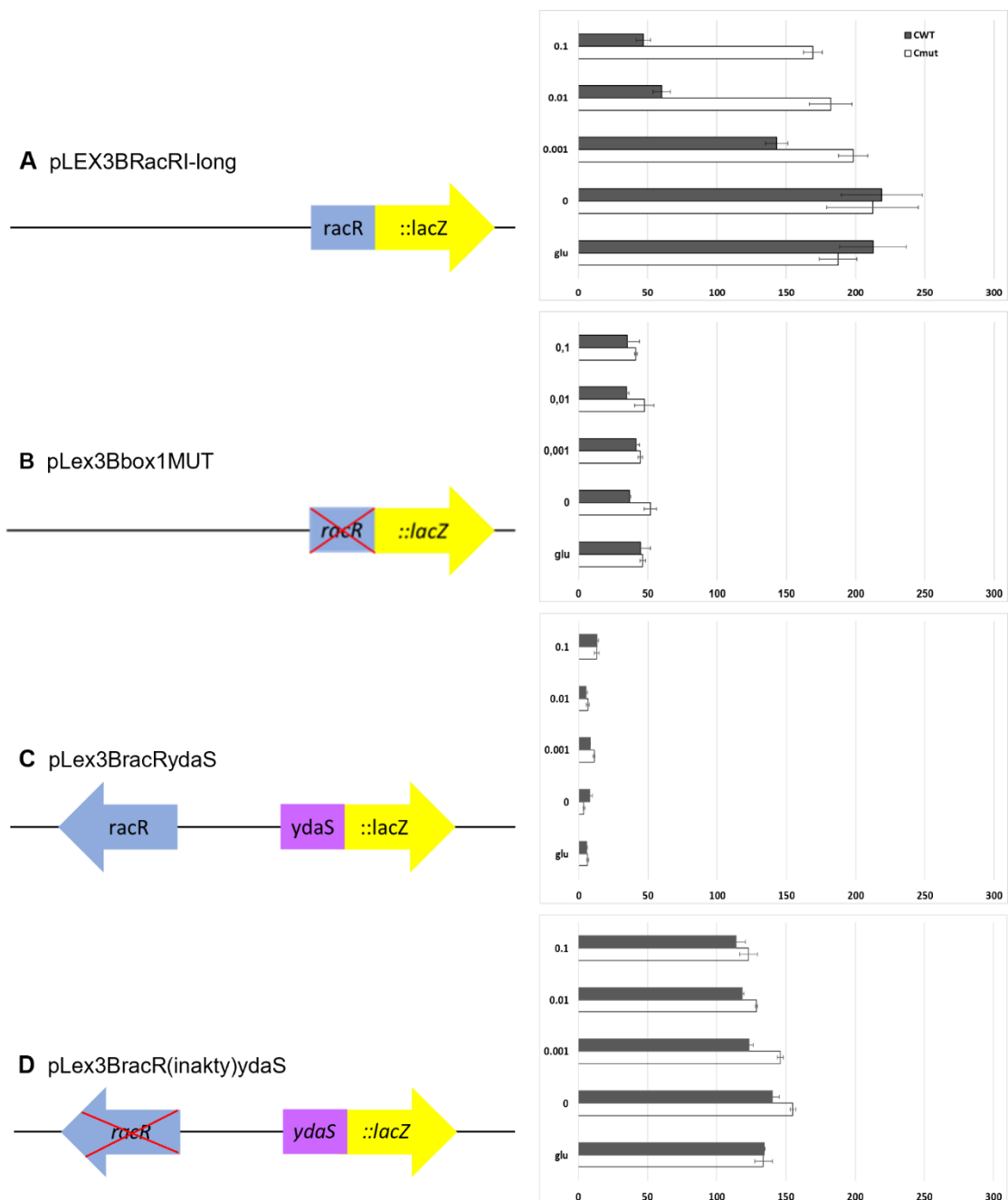


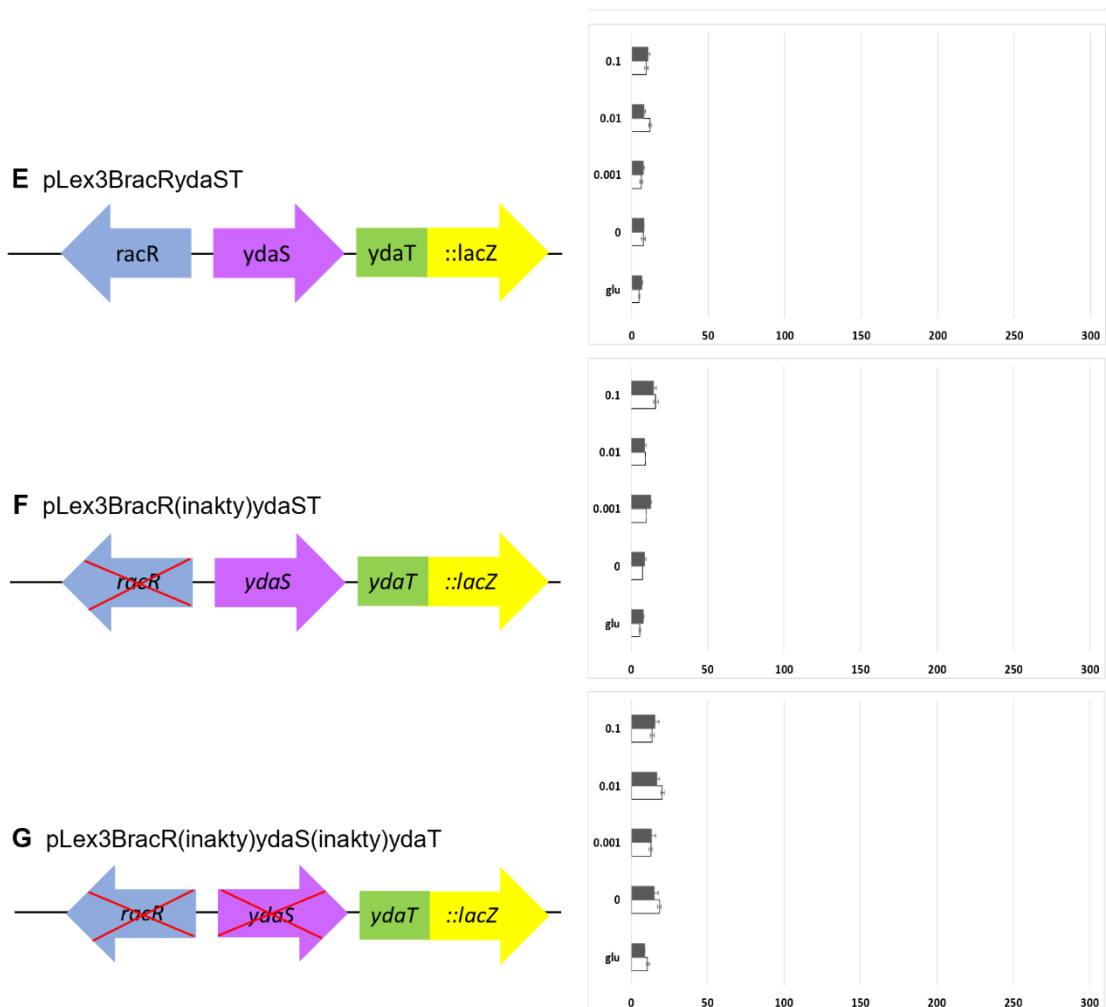
Ryc. 28. Schemat serii konstruktów wykorzystanych do modelu doświadczalnego badającego wpływ białka C na aktywność transkrypcyjną genów *racR*, *ydaS* i *ydaT*.

Zatem mój model doświadczalny zawierał dwa kompatybilne plazmidy, pierwszy – posiadający poszczególny wariant testowanych genów w fuzji translacyjnej z genem

lacZ oraz pBAD-CWT lub pBAD-Cmut dostarczający gen C typu dzikiego (WT) lub mutanta genu C, pod indukowalnym promotorem arabinozowym.

Otrzymane transformanty hodowałam w pożywce minimalnej M9 (Materiały 3.6.3.) wykorzystując glicerol jako źródło węgla, a glukozę i różne stężenia arabinozy do uzyskania gradientu stężenia białka regulatorowego C: 0; 0,001%; 0,01%; 0,1%. Poziom aktywności transkrypcyjnej mierzyłam w teście ONPG, zaś aktywność LacZ wyrażałam w jednostkach Millera (Metody 4.19.).





Ryc. 29. Schematyczne przedstawienie aktywności transkrypcyjnej poszczególnych genów *racR*, *ydaS* oraz *ydaT* wraz z ich nazwami (lewa strona). Układ genetyczny zawierał dwa kompatybilne plazmidy: wariant genów *racR*, *ydaS* oraz *ydaT* oraz pBAD-CWT lub pBAD-Cmut. Wpływ białka C (szare słupki) na aktywność transkrypcyjną poszczególnych genów badałam w obecności wzrastającej ilości białka C. W komórce wykorzystywałam promotor arabinozowy: od braku C (0,1% glukoza) do akumulacji C (0,1% arabinoza). Równolegle kontrolnie badałam wpływ wariantu Cmut (białe słupki). Aktywność transkrypcyjną sprawdzanego genu sprawdzałam w jego fuzji z genem reporterowym *lacZ* mierzalnym ilościowo w teście Millera z ONPG (prawa strona) (Metoda 4.19.).

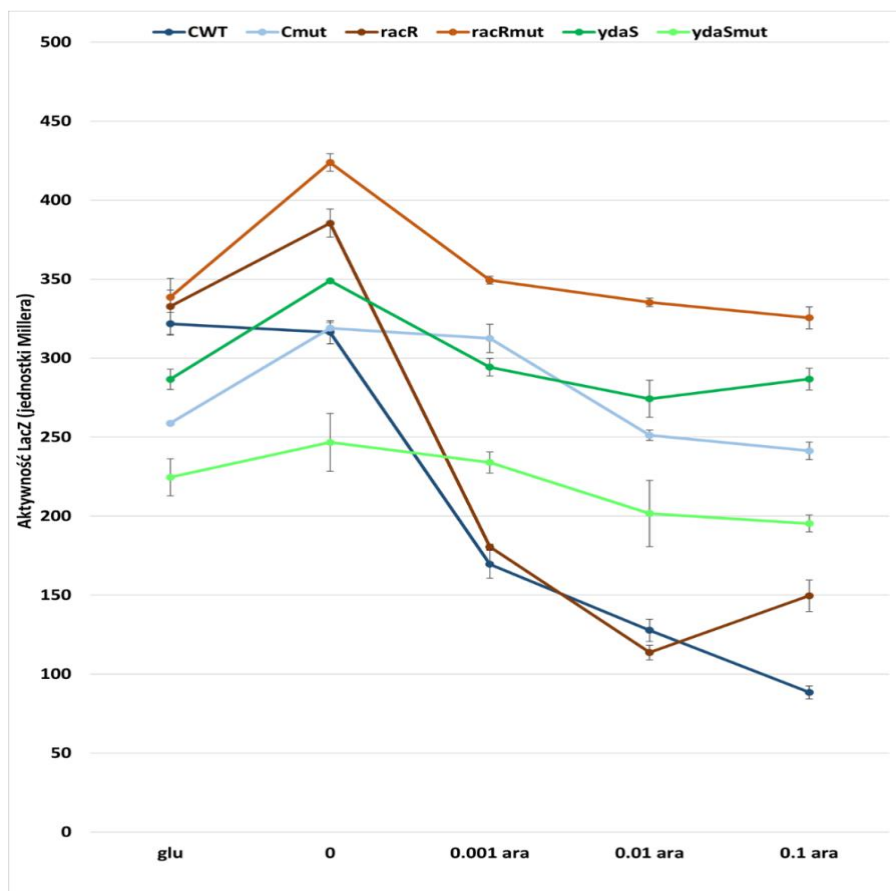
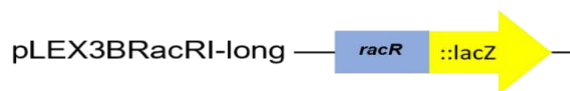
Reasumując wyniki z ryciny 29, z badań ekspresji poszczególnych genów wynika, że:

- 1) ekspresja genu *ydaT* w badanych warunkach nie zachodzi lub jest zablokowana,
- 2) ekspresja genu *racR* wpływa negatywnie na aktywność transkrypcyjną genu *ydaS*, która jest i tak na dość niskim poziomie,
- 3) ekspresja genu CWT wpływa negatywnie na ekspresję genu *racR*,
- 4) wydaje się, że tylko gen *racR* (a nie *ydaS*, ani *ydaT*) jest wrażliwy na efekt białka C typu dzikiego, zatem możliwa jest interakcja białka C z rejonem DNA kontrolującym ekspresję *RacR*.

5.10. Czynniki transkrypcyjne i ich wpływ na ekspresję genów operonu *RacR*

Wydaje się z literatury, że *RacR* jest czynnikiem transkrypcyjnym (Bindal et al., 2017; Krishnamurthi et al., 2017), ale z naszych analiz bionformatycznych wynika, że także *YdaS* oraz *YdaT* mogą być również czynnikami transkrypcyjnymi. Wszystkie te białka są relatywnie małe i posiadają motywy wiązania z DNA. Postanowiłam zbadać, jak te przypuszczalne czynniki transkrypcyjne oraz regulator C wpływają na ekspresję genu *racR* w teście *in vivo*. W doświadczeniu tym wykorzystałam szczep *E. coli* MG1655 Δ *rac*, który zawierał dwa kompatybilne plazmidy: jeden z genem *racR* w fuzji z genem reporterowym *lacZ*. Drugi plazmid niósł czynnik transkrypcyjny (*RacR*, *YdaS* lub gen C) typu dzikiego lub zmutowany (kontrolnie) pod indukowalnym promotorem arabinozowym. Profile transkrypcyjne mierzyłam za pomocą aktywności *lacZ* w jednostkach Millera przy wzrastających stężeniach TF od 0 (glukoza) do najwyższego stężenia (0,1% arabinozy). Nie udało się sklonować genu *ydaT*, także go nie testowano.

Otrzymane wyniki pokazują, że *racR* hamuje własną transkrypcję z promotora *racR*, zatem jest autorepresorem. Ciekawe, że profil ekspresji genu *racR* w obecności gradientu stężeń represora *RacR* i białka C jest porównywalny, ponieważ białko C jest zdolne do redukcji ekspresji *racR*, niemal w takim samym stopniu jak naturalny represor *RacR*. Ponadto otrzymane wyniki wskazują, że białko *YdaS* nie wpływa na ekspresję genu *racR* w żadnym badanym stężeniu (jedynie niewielka różnica poziomu *LacZ* dla *YdaS* wobec *YdaSmut*), w naszych warunkach (Ryc. 30.).

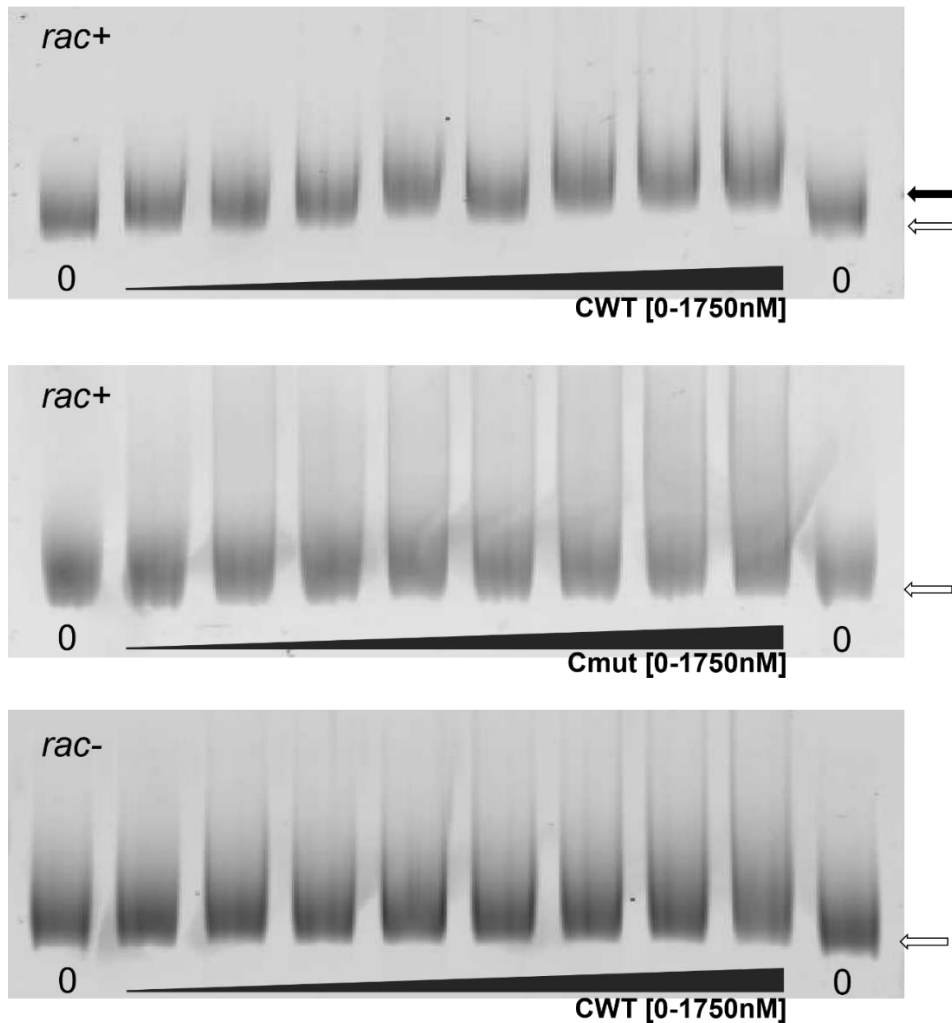


Ryc. 30. Profil transkrypcyjny *racR* mierzony w teście Millera. Każdy model badawczy zawierał dwa plazmidy: jeden z promotorem i genem *racR* w fuzji z genem reporterowym *lacZ* oraz drugi: z badanym TF: C (granatowa linia), C-mut kontrolnie (jasnoniebieska linia), *racR* (brązowa linia), *racR*-mut kontrolnie (pomarańczowa linia), *ydaS* (ciemnozielona linia), *ydaS*-mut kontrolnie (jasnozielona linia). Słupki błędów wskazują średnie odchylenia z trzech niezależnych pomiarów.

5.11. Badanie zdolności wiązania się białka C do rejonu *rac*

W celu sprawdzenia zdolności wiązania się białka C przeprowadziłam test EMSA (ang. *Electrophoretic Mobility Shift Assay*). Chciałam sprawdzić, czy wynik przeprowadzonych eksperymentów *in vivo* był spowodowany przez bezpośrednie wiązanie się białka C do rejonu *rac* w obrębie promotora genu *RacR*. Za pomocą testu EMSA można zbadać interakcję pomiędzy DNA, a białkiem, które potencjalnie rozpoznaje i oddziałuje z sekwencją DNA. Wykorzystuje się zasadę, że powstały kompleks DNA-białko migruje w żelu z mniejszą szybkością niż wolne DNA. Do eksperymentu wykorzystałam dwa substratowe DNA. Jeden krótszy o długości 175 p.z., który obejmował operator genu *racR* i jego sekwencje kodującą, oraz dłuższy o długości

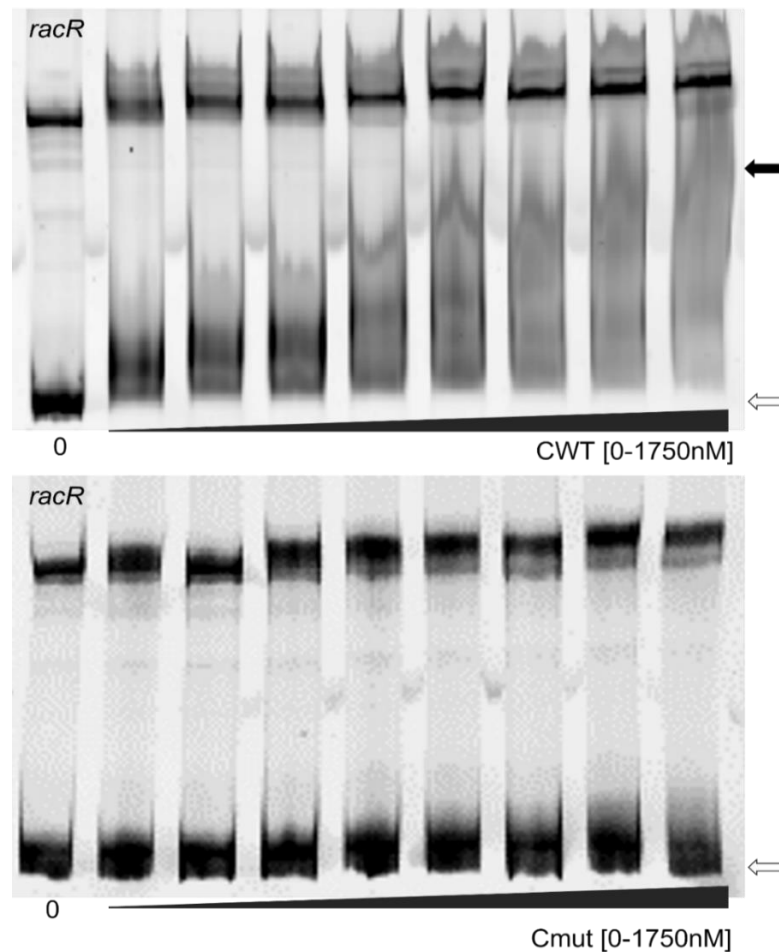
271 p.z., który obejmował cały region międzygenowy (IGR) pomiędzy genami *racR*, a *ydaS* i część sekwencji kodującej *racR* (Materiały 3.3.). DNA amplifikowałam metodą PCR z jednym starterem wprowadzającym znacznik fluorescencyjny Cy5 na końcu 5'. Każdą reakcję wiązania przeprowadzałam z taką samą ilością DNA (25nM) i rosnącym stężeniem białek CWT oraz Cmut jako kontrola negatywna (0; 500; 650; 800; 1000; 1150; 1300; 1500; 1750 nM) (Metoda 4.21.) (Ryc. 31.).



Ryc. 31. Test EMSA sekwencji *rac* wiążącej białko C. DNA substratowy o długości 175 p.z., obejmował operator genu *racR* i jego sekwencje kodującą (oznaczony jako *rac+*). Kontrolnie użyto inny substrat DNA (tzw. sekwencje niespecyficzną - *rac-*) o zbliżonej długości. Czarną strzałką oznaczyłam tworzący się kompleks o wolniej migrujący, natomiast białą – niezwiązane DNA.

Analizując test EMSA, na krótszym fragmencie mogę potwierdzić zmianę drogi migracji DNA-białko tylko dla kombinacji sekwencji *racR+* i białka CWT, ale nie dla innych kontrolnych kombinacji, takich jak *racR+* / Cmut lub *racR-* / CWT (Ryc. 31.).

Natomiast jeżeli chodzi o test EMSA wykonany w identycznych warunkach na dłuższym fragmencie, zmianę migracji DNA-białko również zaobserwowałam tylko dla białka typu dzikiego. Opóźnione kompleksy dla białka typu dzikiego były raczej rozmyte i prawdopodobnie niestabilne, a przy wyższym stężeniu prążek substratu DNA był niewidoczny (Ryc. 32.). Sekwencja DNA genu *racR* nie jest naturalnym celem wiązania dla białka C, dlatego spodziewaliśmy się raczej słabego wiązania.

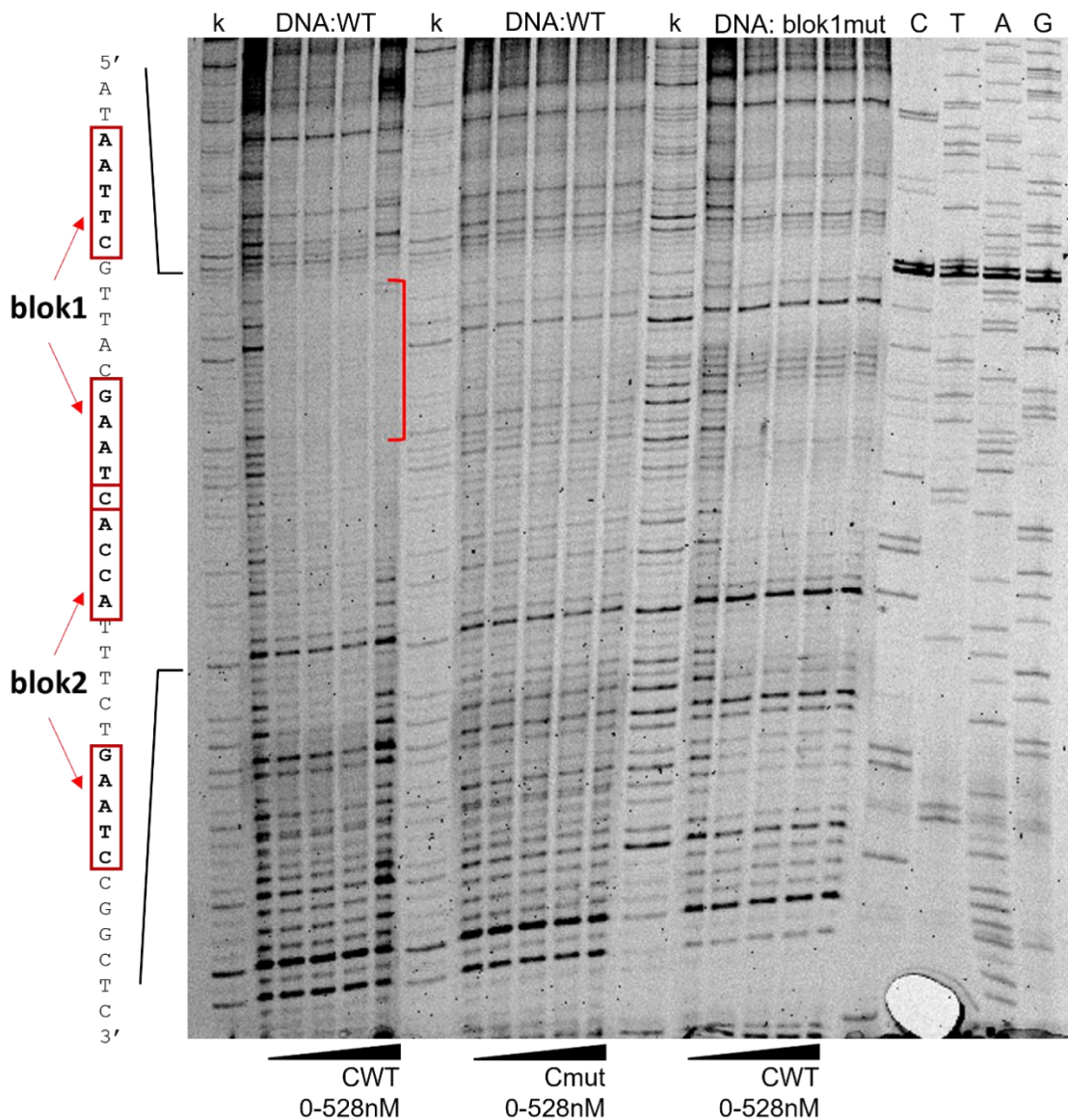


Ryc. 32. Wiązanie się białka C do sekwencji *rac* o długości 271 p.z. zawierający cały rejon intergenowy IGR (górny żel) oraz brak wiązania dla kontrolnego białka Cmut (dolny żel). Czarną strzałką oznaczyłam tworzący się kompleks, natomiast białą – niezwiązane DNA.

5.12. Ustalenie miejsc wiązania białka C do sekwencji nukleotydowej rejonu faga kryptycznego *Rac* z użyciem metody śladu DNA (ang. *DNA footprinting*)

Po przeprowadzonym teście EMSA, który pokazał wiązanie się białka C do sekwencji *rac*, do krótszego (175 p.z.) oraz dłuższego (271 p.z.) substratu DNA (Wyniki 5.11.), postanowiłam zmapować dokładne miejsca wiązania się białka C do tej sekwencji

nukleotydowej. Metodą śladu DNA (ang. *footprinting*) (Metody 4.22.) można ustalić rejon wiązania się białka, niemal co do nukleotydu, który wynika z tego, że DNazaI nie jest w stanie dokonać cięcia DNA, jeśli ten rejon jest w kompleksie z białkiem. W pierwszym etapie za pomocą reakcji PCR (startery RACRplexC, PFracRFLUO1Cy5znacznik (Materiały 3.3.) przygotowałam substratowy DNA o długości 175 p.z., który obejmował sekwencję kodującą *racR* oraz rejon IGR. Powstały produkt wyznakowany znacznikiem fluorescencyjnym Cy5 użyłam jako substrat DNA do przeprowadzenia eksperymentu *in vitro*.



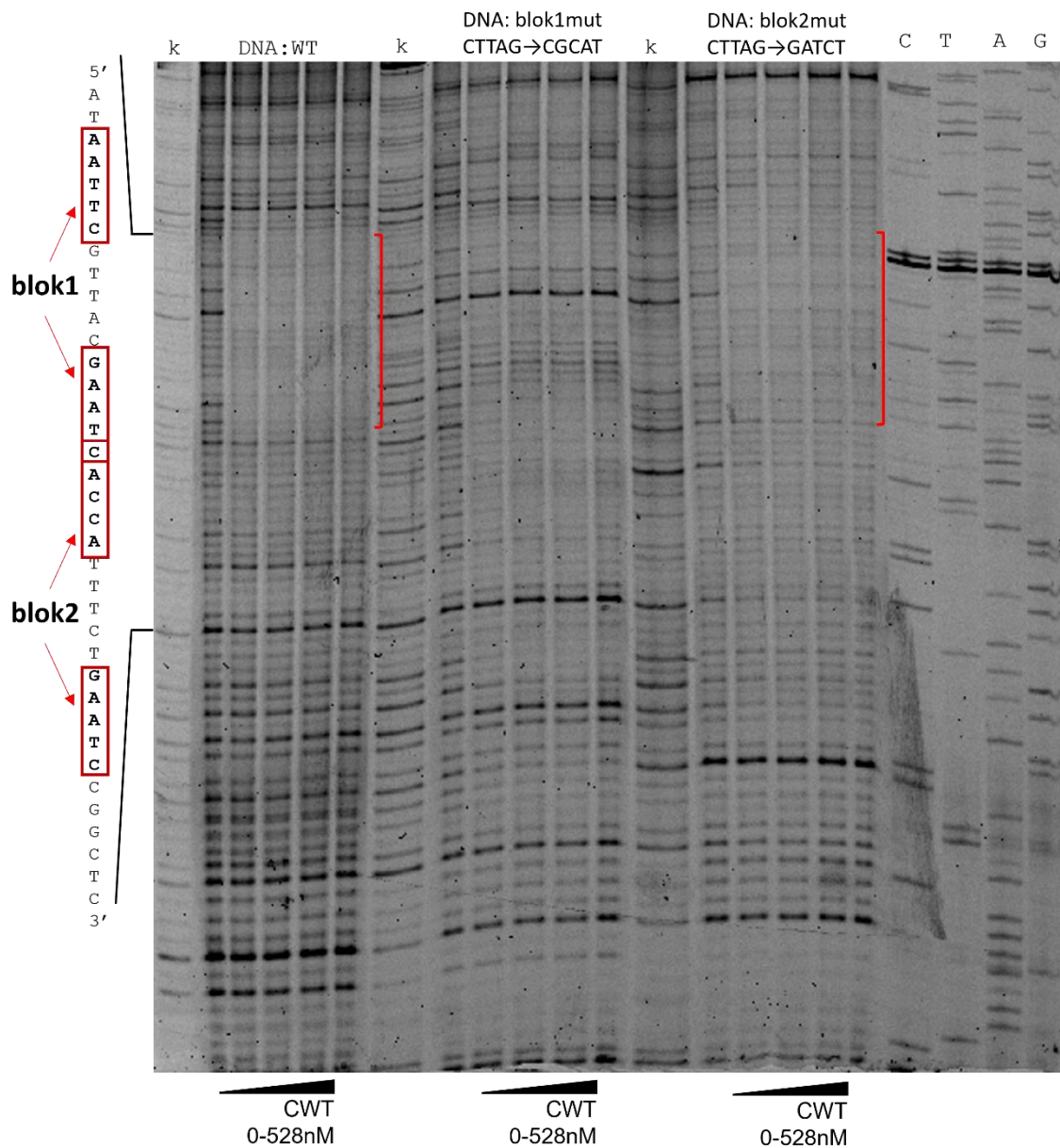
Ryc. 33. Mapowanie miejsca wiązania regulatora C w obrębie rejonu faga *rac* z użyciem testu *footprint*. Czerwoną kłmramą oznaczyłam miejsca chronione przed cięciem DNazaI. Użyto wzrastających stężeń białka CWT oraz Cmut oraz dwóch substratów DNA, typu dzikiego WT oraz DNA z mutacją w obrębie tzw. „bloku1” (5'-CTTAG-3' zmienionego na 5'-CGCAT-3'). Ścieżki oznaczone k stanowią kontrolę – substrat DNA wyłącznie poddany działaniu DNazyI.

Przygotowany substratowy DNA był identycznej długości, jak ten w reakcji EMSA. Przeprowadziłam test *footprint* przy użyciu białka C i jego wariantu Cmut jako kontroli negatywnej (Metody 4.22.). Zaobserwowałam wyraźny region ochrony przed cięciem DNazyI, którego nie było na żelu w kontroli z użyciem białka Cmut (Ryc. 33.). Sekwencja nukleotydowa tego rejonu została przeanalizowana pod kątem podobieństwa do tzw. sekwencji consensus wyznaczonej dla regulatora C: 5'-CTAAG-n5-CTTAG-3', które tworzą odwrócone powtórzenia rozdzielone 5nt łącznikiem (Rezulak et al., 2016). Znaleźliśmy dwa miejsca 5'-CTAAG-3' na komplementarnej nici DNA wobec nici kodującej gen *racR*, położone blisko startu translacji genu *racR*. Sekwencje te mogą tworzyć dwie alternatywne, potencjalne sekwencje wiązania białka C, tj. sekwencja oznaczona jako „blok1” (5'-CTAAG-n5-CTTAA-3') oraz „blok2” (5'-CTAAG-n5-ACCAC-3') (Ryc. 33.). Sekwencja „bloku1” różniła się tylko 1nt od sekwencji consensus białka C, ale z uwagi na ich bliskość postanowiliśmy zweryfikować, czy faktycznie tylko „blok1” jest chroniony przed cięciem DNaząI. Przygotowaliśmy substrat DNA, który ma zmutowaną sekwencję „bloku1”, tj. 5'-CTTAG-3' → 5'-CGCAT-3'. Zauważyłam ostre prążki odpowiadające produktom trawienia przez DNazęI dla fragmentu DNA zmutowanego, w obrębie rejonu, który uprzednio był chroniony przed trawieniem (Ryc. 33 i 34).

Mimo to, postanowiliśmy także wykluczyć możliwość wiązania się do obu sekwencji „bloku1” oraz „bloku2”. Przygotowaliśmy jeszcze jeden substrat DNA, tym razem ze zmienioną sekwencją w obrębie „bloku2” (5'-CTTAG-3' → 5'-CGCAT-3') i powtórzyliśmy procedurę (Ryc. 34.). Wyniki pokazują, że taki zmieniony DNA wydaje się być chroniony w rozpatrywanym rejonie, w odróżnieniu od substratu ze zmutowanym „blokiem1”.

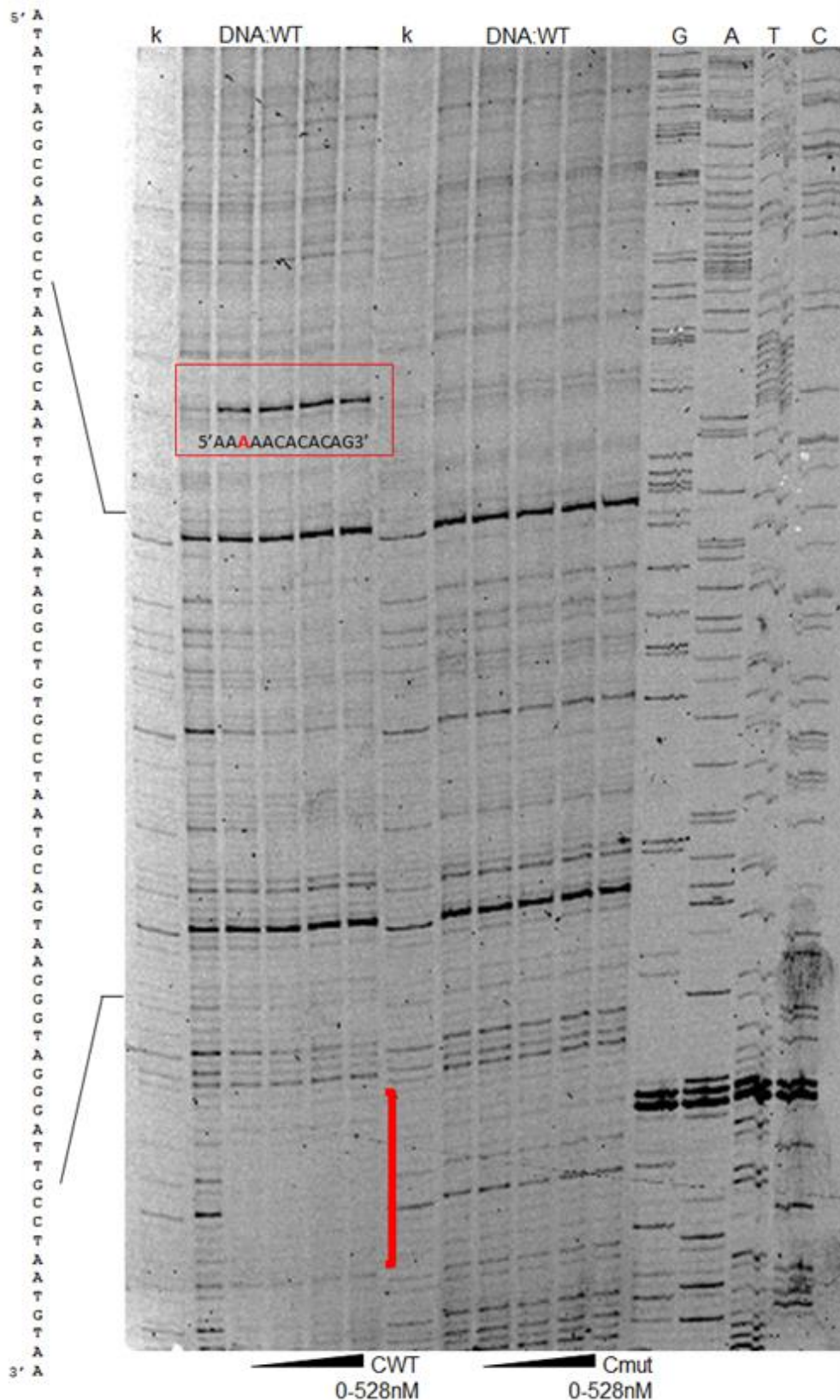
Reasumując wyniki mapowania, wykazałam, że białko C wiąże się do sekwencji niepowiązanej z funkcją regulatora C, nie naturalnej i być może to wiązanie jest „przypadkowe” z punktu widzenia fizjologii komórki. Niemniej jednak położenie tej sekwencji (5'-CTAAG-n5-CTTAA-3') w dużej bliskości od kodonu startu translacji genu *racR* może wywierać negatywny wpływ na ekspresję represora RacR i jego regulonu.

Dodatkowo kontrolnie wykonałam również *footprint* w regionie intergenowym (IGR) powyżej promotora genu *racR*, aby wykluczyć tam możliwe dodatkowe miejsca wiązania. Nie znalazłam żadnego dodatkowego obszaru chronionego przed DNaząI poza już odkrytym obszarem w obrębie „bloku1” (Ryc. 35.).



Ryc. 34. Mapowanie miejsca wiązania regulatora C w obrębie rejonu faga *rac* z użyciem testu *footprint*. Czerwoną klamrą oznaczyłam miejsca chronione przed cięciem DNaząI. Użyto wzrastających stężeń białka CWT oraz trzech substratów DNA, typu dzikiego WT oraz DNA z mutacją w obrębie tzw. „bloku1” lub „bloku2” (5'-CTTAG-3' zmienionego na 5'-CGCAT-3'). Ścieżki oznaczone k stanowią kontrolę – substrat DNA wyłącznie poddany działaniu DNazyI.

Jednak analizy uzyskanego żelu ujawniły obecność sekwencji, które były „lepszym” substratem dla trawienia przez DNazęI (tzw. „hypersensitive site”), co zwykle świadczy o tym, że białko C może silnie zginać DNA, co ułatwia dostępność DNazyI w innych regionach substratu. W tym wypadku region ten jest bogaty w adeniny (5'-AAAACACACA-3'), położony jest w obrębie startu transkrypcji genu *ydaS* (Ryc. 35.).



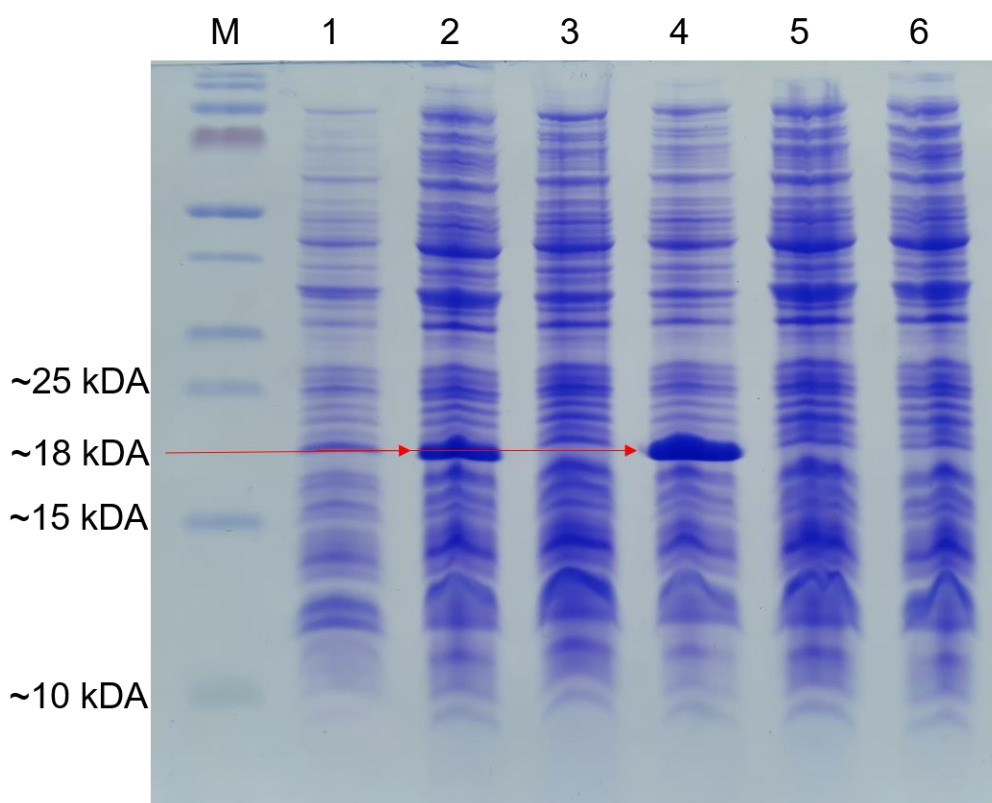
Ryc. 35. Test śladu DNA (ang. *footprinting*) na matrycy DNA rejonu intergenowego faga *rac*. Czerwoną kłmą oznaczyłam widoczne miejsca protekcji przed cięciem DNazyI. Ramką oznaczyłam występowanie miejsc tzw. „hypersensitive site”, gdzie sekwencja DNA była bardziej wrażliwa na cięcie DNazyI. Ścieżki oznaczone k stanowią kontrolę – substrat DNA wyłącznie poddany działaniu DNazyI.

5.13. Oddziaływanie białka RacR z rejonem *rac*

Do dalszych badań nad mechanizmem molekularnym wpływu białka C na regulon regulatora RacR potrzebny był oczyszczony preparat białka RacR.

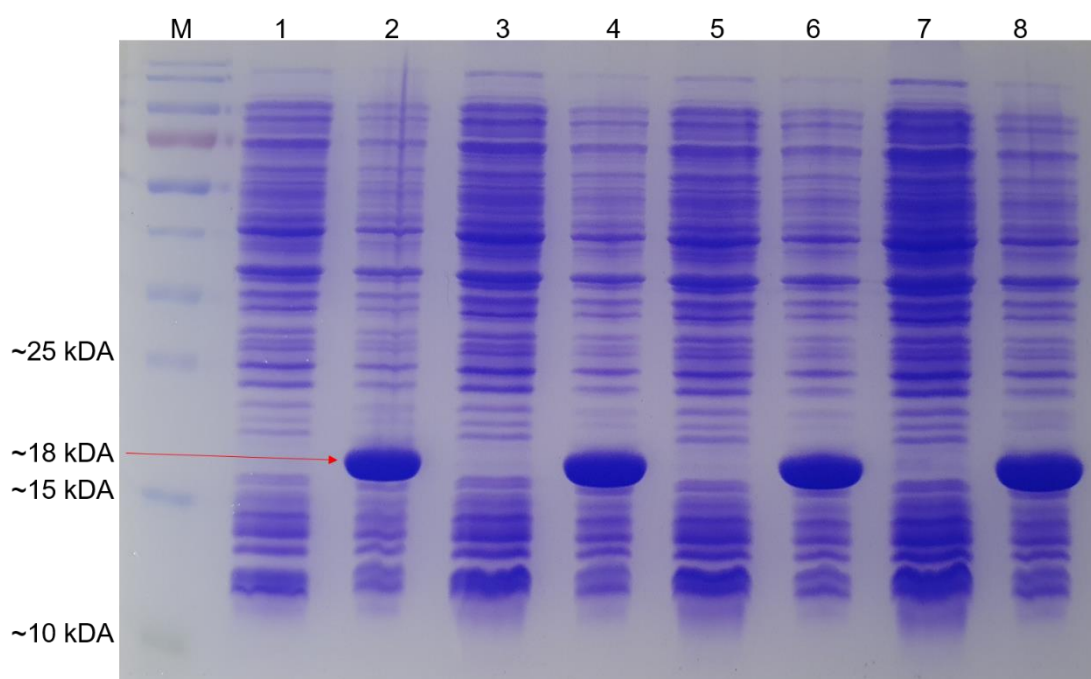
5.13.1. Nadprodukcja i oczyszczanie białka RacR

Do badania oddziaływań białka RacR z rejonem *rac* potrzebowałam oczyścić to białko. W celu nadprodukcji białka RacR skonstruowałam plazmid pET28a(+)*racR* (Materiały 3.2.), z wklonowanym genem *racR* z sekwencją znacznika histydynowego (His-6) w którym ekspresja genu *racR* odbywa się w układzie ekspresyjnym T7 (Tabor & Richardson, 1985). Próby nadprodukcji RacR wykonałam w komórkach *E. coli* BL21(DE3), ER2566 oraz Rosetta(DE3) (Materiały 3.1.). Najbardziej wydajna nadprodukcja zaszła w komórkach *E. coli* BL21(DE3) i ER2566 (Rys. 36.). Z tego powodu wybrałam komórki *E. coli* BL21(DE3) do oczyszczania białka RacR.



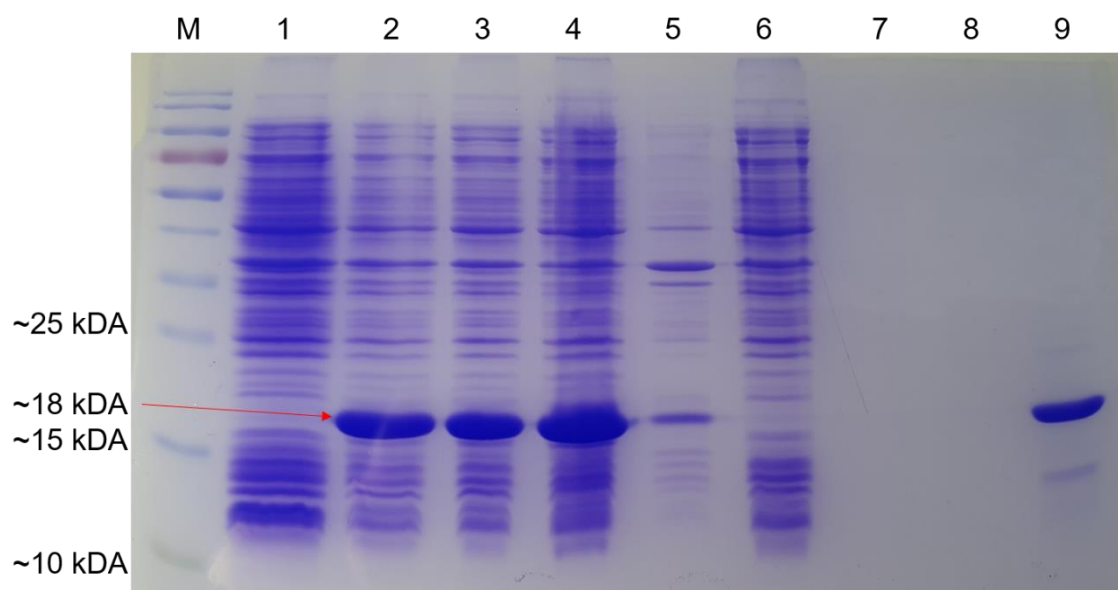
Ryc. 36. Rozdział elektroforetyczny lizatów komórkowych po nadprodukcji białka RacR na 12,5% żelu poliakrylamidowym. 1. Próbką przed indukcją w *E. coli* BL21(DE3), 2. Próbką po indukcji IPTG w *E. coli* BL21(DE3), 3. Próbką przed indukcją w *E. coli* ER2566, 4. Próbką po indukcji IPTG w *E. coli* ER2566, 5. Próbką przed indukcją w *E. coli* Rosetta(DE3), 6. Próbką po indukcji IPTG w *E. coli* Rosetta(DE3). M – wskaźnik masowy białek PageRuler Prestained Protein Ladder. Czerwone strzałki oznaczają pozycje nadprodukowanych białek.

Nadprodukowane rekombinowane białka RacR z komórek *E. coli* BL21(DE3) poddałam oczyszczaniu metodą chromatografii powinowactwa (Metody 4.13.) na kolumnie TALON, która to dzięki jonom kobaltowym umożliwia wiązanie znaczników His-tag czyszczonych białek. Z kolejnych etapów oczyszczania pobierałam próbki, które następnie rozdzielałam na żelu poliakrylamidowym w celu sprawdzenia obecności białka (Ryc. 37.).

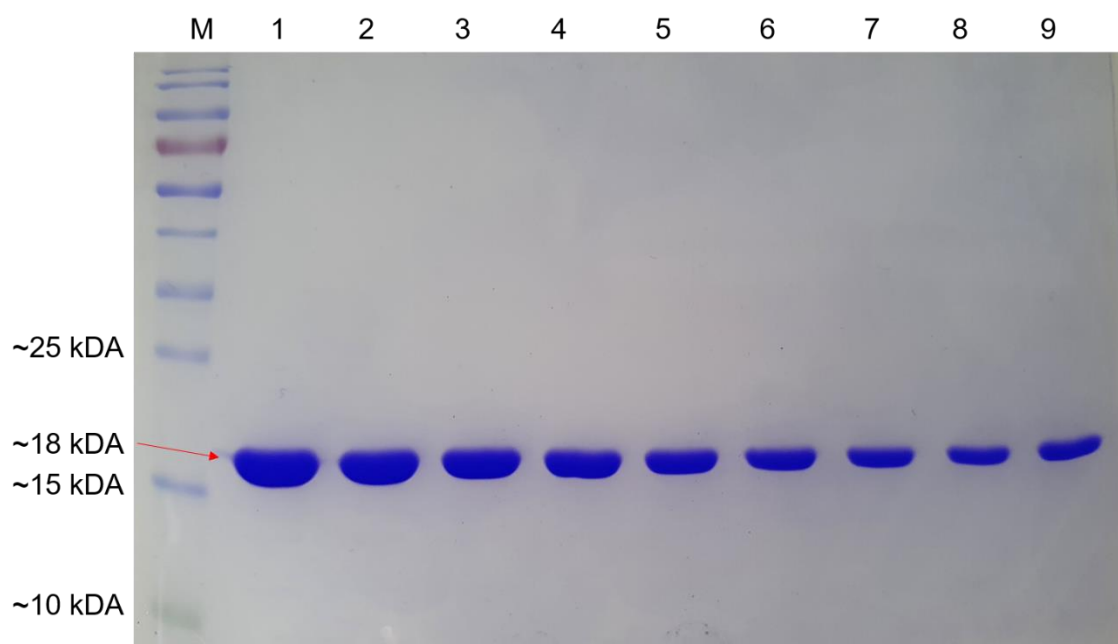


Ryc. 37. Rozdział elektroforetyczny preparatów białkowych po nadprodukcji białka RacR na dużą skalę w komórkach *E. coli* BL21(DE3) na 12,5% żelu poliakrylamidowym. 1,3,5,7 - próbki przed indukcją w *E. coli* BL21(DE3), 2,4,6,8 - próbki po indukcji IPTG w *E. coli* BL21(DE3). M – wskaźnik masowy białek PageRuler Prestained Protein Ladder. Czerwona strzałka oznacza pozycje nadprodukowanego białka.

Po uzyskaniu osadów rekombinowanego białka poddałam je lizie komórkowej wykorzystując ultradźwięki - sonikację i odpowiednie etapy oczyszczania białka. Na każdym etapie oczyszczania pobierałam próbki oraz dokonywałam rozdziału próbek w 12,5% żelu poliakrylamidowym (Ryc. 38, 39).



Ryc. 38. Rozdział preparatów białkowych z poszczególnych etapów oczyszczania RacR na 12,5% żelu poliakrylamidowym. Kolejno zostały umieszczone próbki: 1. Kontrola bez nadprodukcji, 2. Nadprodukowane białko, 3. Supernatant po sonifikacji, 4. Supernatant po wirowaniu, 5. Osad po wirowaniu, 6. Przesącz, 7. Przemycie buforem NPi-10, 8. Przemycie buforem NPi-20, 9. Pierwsza elucja białka. M - wzorzec masowy białek PageRuler Prestained Protein Ladder. Czerwona strzałka oznacza pozycje nadprodukowanego białka.

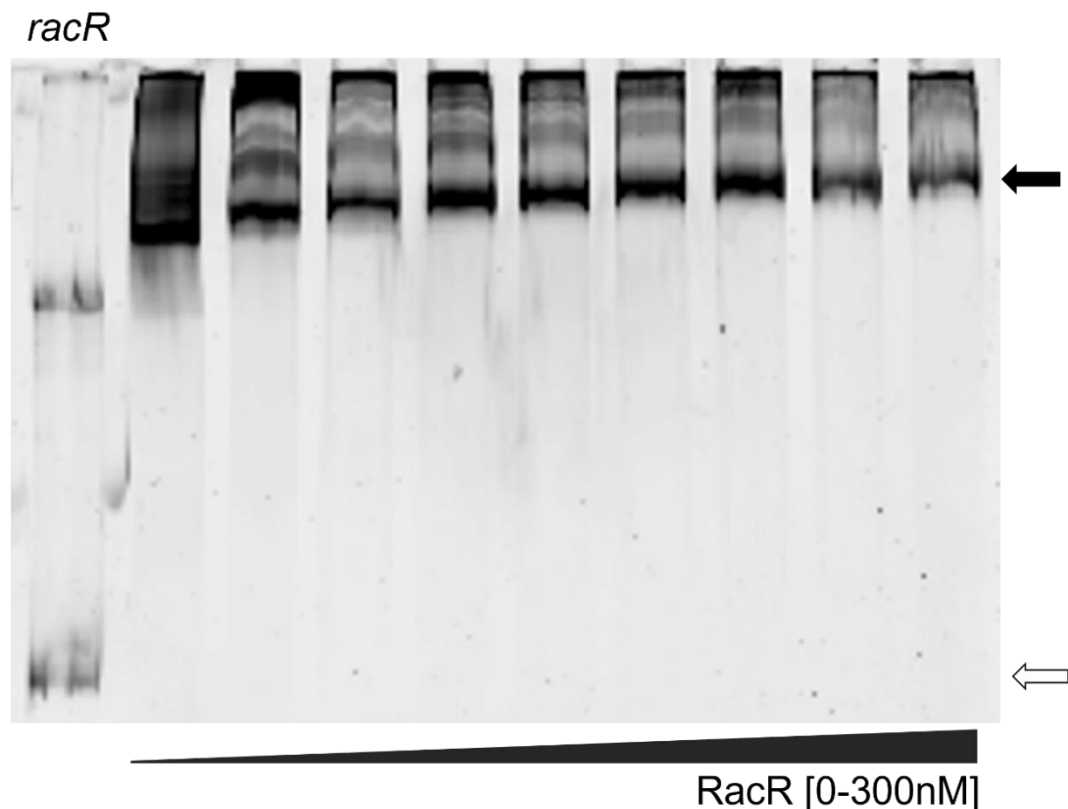


Ryc. 39. Rozdział preparatów białkowych z elucji białka RacR na 12,5% żelu poliakrylamidowym. Kolejno zostały umieszczone próbki: 1-9 elucji białka. M - wzorzec masowy białek PageRuler Prestained Protein Ladder. Czerwona strzałka oznacza pozycje nadprodukowanego białka.

Ostatecznie uzyskałam 4 ml preparatu RacR o stężeniu 0,5 mg/ml.

5.13.2. Badanie interakcji białka RacR z rejonem faga *rac*

W kolejnym kroku postanowiłam przeprowadzić test EMSA dla białka RacR, żeby zweryfikować wiązanie się RacR do jego naturalnego miejsca wiązania w obrębie promotora. Do eksperymentu wykorzystałam substratowy DNA o długości 271 p.z. (ten sam, co w przypadku eksperymentu EMSA dla CWT), który obejmował cały region międzygenowy (IGR) i część sekwencji kodującej *racR* (Materiały 3.3.). Każdą reakcję wiązania przeprowadzałam z taką samą ilością DNA (25nM) i rosnącym stężeniem białka RacR (0; 27; 54; 81; 108; 135; 162; 189; 243; 300 nM). Jak można zauważyć w przeciwieństwie do wiązania się białka C do tego rejonu (Ryc. 31, 32) represor RacR wiąże się do badanej sekwencji DNA przy bardzo niskim stężeniu tego białka – już przy 33 nM. Ponadto eksperyment wykazywał bardzo wyraźnie tworzące się wolno migrujące kompleksy DNA-RacR, co wskazuje na silne i specyficzne wiązanie oraz możliwość dużej oligomeryzacji białka na substracie DNA (Ryc. 40.).

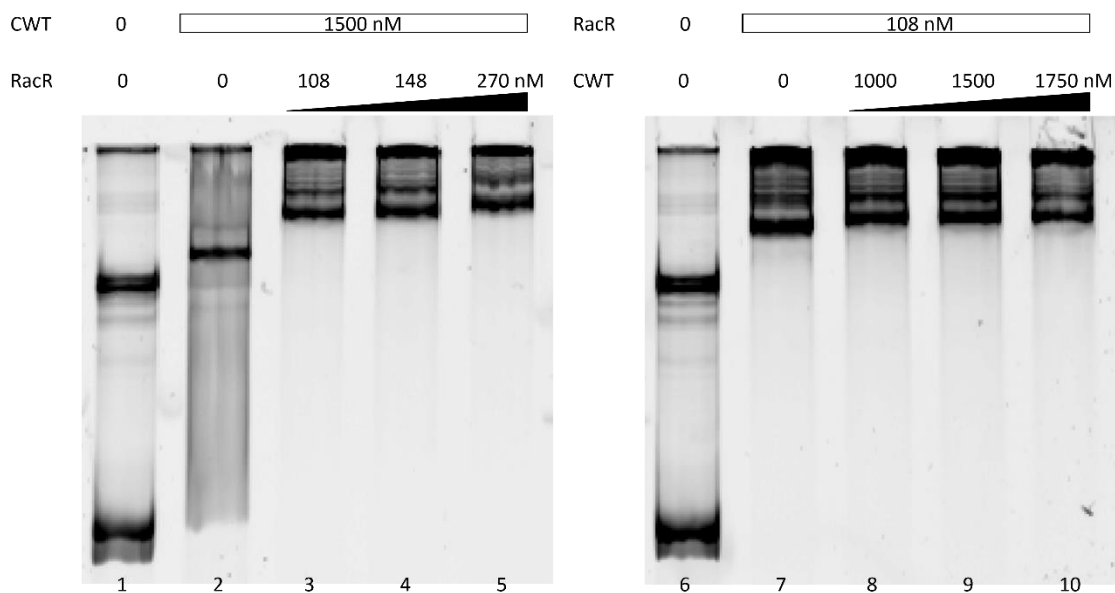


Ryc. 40. Wiązanie się białka RacR do sekwencji faga *rac*. Czarną strzałką oznaczyłam tworzący się wolno migrujący kompleks DNA-białko, natomiast białą – niezwiązane DNA.

5.14. Oddziaływanie kompetycyjne białek C i RacR w obrębie rejonu faga rac

W następnej kolejności postanowiłam przeprowadzić reakcje EMSA z dwoma czynnikami transkrypcyjnymi: białkiem C i represorem RacR w interakcji z substratem DNA, by zweryfikować, czy białka konkurują o miejsce wiązania, czy może w naszych testowanych warunkach *in vitro* oba wiązania są możliwe, bez wypierania jednego z nich. W celu wykrycia możliwej interferencji w wiązaniu między tymi dwoma represorami wykonałam tzw. konkurencyjny EMSA. Do eksperymentu wykorzystałam ten sam substratowy DNA o długości 271 p.z., który obejmował cały region międzygenowy (IGR) i część sekwencji kodującej *racR* (Materiały 3.3.). Na początku dodałam do czterech reakcji 1500 nM białka C i inkubowałam przez 20 minut w temperaturze pokojowej, aby utworzyć stabilny kompleks z DNA (reakcja 2, Ryc. 41.). Po upływie tego czasu dodawałam wzrastające stężenie białka RacR (0-270 nM), które również inkubowałam taki sam czas w takiej samej temperaturze. Na Rycinie 41 (lewy panel, reakcje 3-5) można zaobserwować, że pomimo stabilnego związanie się białka C, uzyskano kompleksy wyższego rzędu po dodaniu RacR (3-5 studzienka) i to kompleksy wolniej migrujące niż sam kompleks DNA-RacR (reakcja 7). Następnie postanowiłam sprawdzić odwrotną sytuację, dlatego dodałam najpierw 108 nM białka RacR do czterech reakcji (7-10 studzienka) i inkubowałam przez 20 minut w temperaturze pokojowej, po upływie tego czasu dodawałam wzrastające stężenie białka C (0-1750 nM do reakcji 8-10), które również inkubowałam taki sam czas w takiej samej temperaturze. Wydaje się, że powstają kompleksy jeszcze wyższego rzędu w reakcjach z oboma białkami, ale są tak wolno migrujące, że rozdzielczość żelu wydaje się zbyt mała, aby wyciągnąć jednoznaczną konkluzję. Możliwe, że białka wiążą się do tego samego rejonu (lub blisko położonego) DNA, ale nie są w stanie zepchnąć już związanego białka. Możliwe też, że różnią się powinowactwem do wiązania z DNA lub białko RacR ma wiele miejsc wiązania, w odróżnieniu od pojedynczego miejsca wiązania regulatora C określonego w sekwencji sąsiadującej ze startem translacji genu *racR*.

Z tego powodu postanowiono, dokładnie określić miejsce wiązania represora RacR podejmując podejście eksperymentalne wykorzystane wcześniej, czyli reakcje typu test śladu DNA (ang. *footprint*).

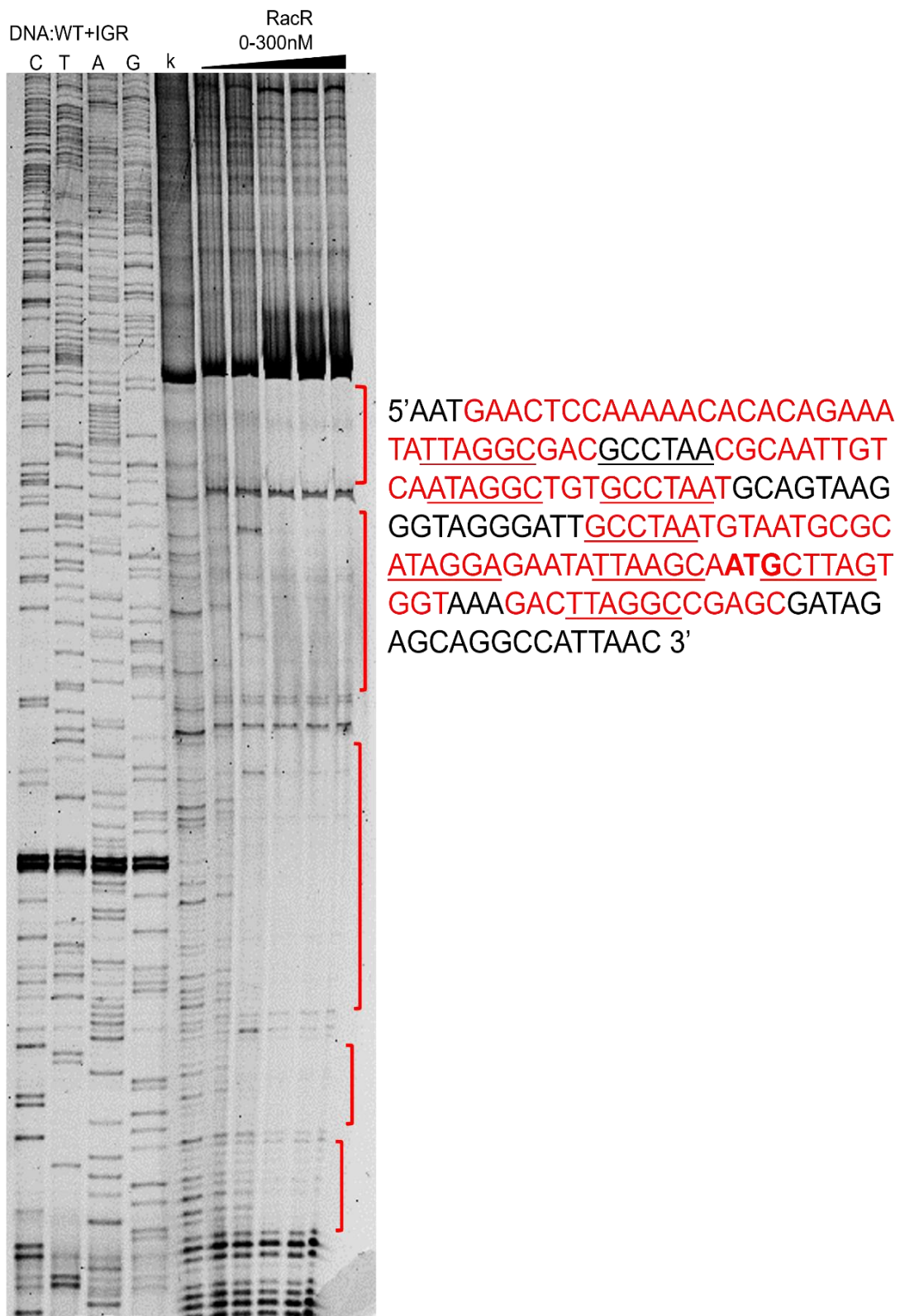


Ryc. 41. Test konkurencyjnego wiązania się białek: regulatora C oraz represora RacR do tego samego substratu DNA. Lewy panel: w warunkach ustabilizowanego wiązania się białka C (20min), dodawano wzrastające stężenie RacR, jak zaznaczono na rycinie. Prawy panel: najpierw dodawano represor RacR w stężeniu wysycającym wiązanie się DNA, a po inkubacji 20 minutowej dodawano wzrastające stężenie regulatora C.

5.15. Ustalenie miejsca wiązania białka RacR do sekwencji nukleotydu rejonu faga *rac* zawierającej rejon intergenowy

W kolejnym teście chciałam ustalić miejsce wiązania białka RacR do sekwencji nukleotydu rejonu faga *rac*, który obejmował cały region międzygenowy (IGR) i część sekwencji kodującej *racR* (Materiały 3.3.), identyczny do tego wcześniej używanego, także w reakcjach EMSA.

Tym razem test śladu DNA (ang. *footprinting*) z białkiem RacR pokazał duże rejony protekcji przed cięciem DNaząI rozdzielone miejscami wyraźnego cięcia (Ryc. 42.), co świadczy o tym, że w badanym obszarze może znajdować się kilka miejsc wiązania DNA. Pobieżna analiza sekwencji DNA wskazała na obecność wielu odwróconych powtórzeń, zatem w kolejny etapie starałam się skorelować pozycje występowania odwróconych powtórzeń z rejonem protekcji przed cięciem DNaząI (sekwencje na czerwono, Ryc. 42.), by wyodrębnić potencjalny motyw wiązania się RacR do DNA.

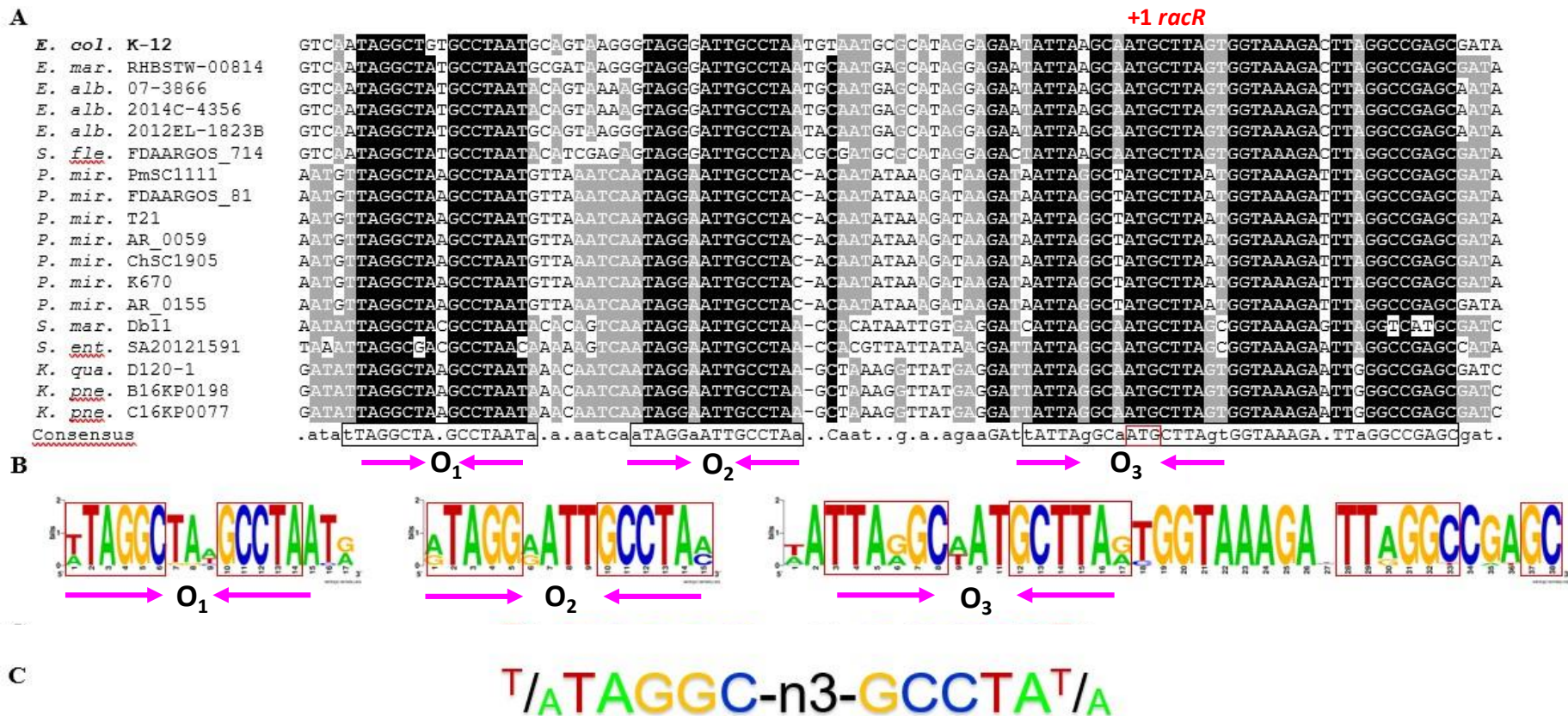


Ryc. 42. Test śladu DNA (ang. *footprinting*) na matrycy rejonu faga *rac* wraz z rejonem intergenowym z użyciem białka RacR. Czerwoną klamrą na żelu oznaczyłam widoczne miejsca wiązania się białka oraz czerwonym kolorem na sekwencji DNA. Znalazłam dziewięć domniemyanych miejsc wiązania represora RacR, które oznaczyłam podkreśleniem w sekwencji DNA. W sekwencji pogrubiałam kodon start dla *racR*. Ścieżki oznaczone k stanowią kontrolę – substrat DNA wyłącznie poddany działaniu DNazyI.

5.16. Analiza *in silico* regionu intergenowego pomiędzy genem *racR*, a *ydaS*

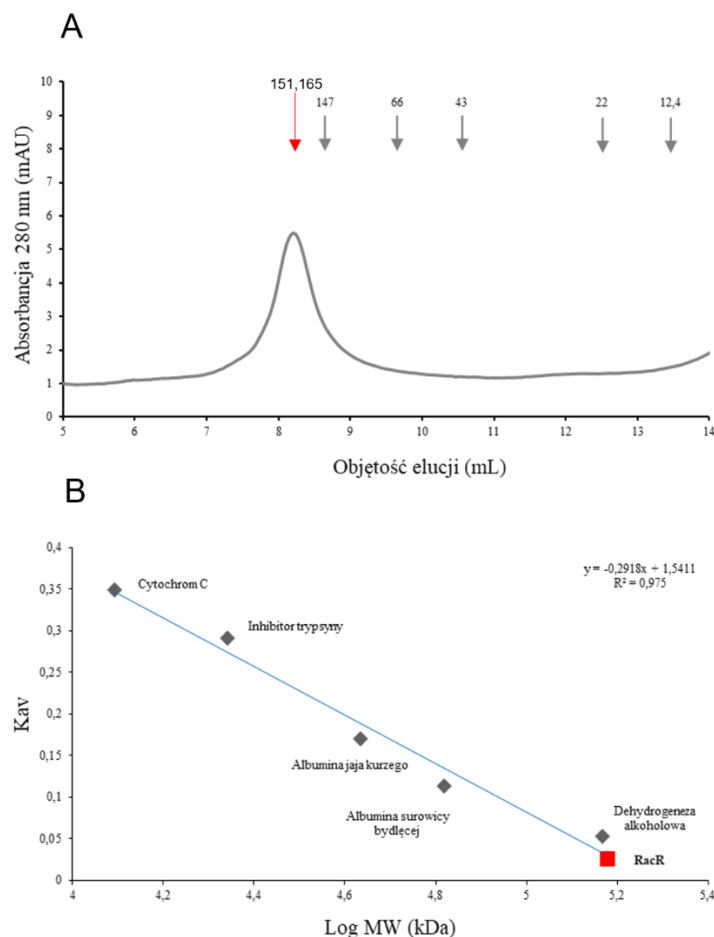
Kolejno, postanowiłam przeprowadzić analizę porównującą sekwencje nukleotydowe w rejonie intergenowym pomiędzy genem *racR*, a *ydaS* dla około 18 różnych gatunków w rodzinie pałeczek jelitowych, dla których stwierdzono występowanie genu *racR*, w tym: *Escherichia marmotae*, *Escherichia albertii*, *Shigella flexneri*, *Proteus mirabilis*, *Salmonella enterica*, *Klebsiella quasipneumoniae* oraz *Klebsiella pneumoniae* (Ryc. 43. A). Analiza ta miała za zadanie pomóc określić motyw wiązania RacR. Znalazłam kilka sekwencji palindromowych, których lokalizacja pokrywa się z eksperymentalnie wyznaczonymi miejscami wiązania się RacR do własnego promotora. Z tych sekwencji postanowiłam stworzyć LOGO, które określa motyw DNA, do którego wiąże się represor RacR (Metody 4.24.) (Ryc. 43. B). Sekwencje te, jak poszczególne wyznaczone przeze mnie operatory, mogą być wiązane z różnym powinowactwem, ponieważ zawierają pojedyncze zmiany w nukleotydach. Zaproponowałam ogólny proponowany motyw wiązania białka RacR do sekwencji intergenowej w postaci sekwencji 5'-^{T/A}TAGGC-n3-GCCTA^{T/A}-3' (Ryc. 43. C). Ten motyw wystąpił w czterech miejscach w badanym substracie DNA, pokrywając region pomiędzy genami *ydaS* i *racR* (trzy operatory) oraz jest obecny w sekwencji kodującej *racR*. Ta czwarta pozycja wiązania RacR nie została oznaczona jako operator, bo nie zawiera kompletnego odwróconego powtórzenia. Wydaje się, że odwrócone powtórzenia mogą być wiązane przez homodimer represora RacR.

Ciekawe, że właśnie pozycja czwarta wiązania RacR pokrywa miejsce wyznaczone dla wiązania się białka C (Ryc. 43. A).



Ryc. 43. A – Porównanie sekwencji nukleotydujowej rejonu intergenowego pomiędzy genem *racR*, a *ydaS* oraz początek startu transkrypcji dla genu *racR* oznaczony czerwoną ramką. Czarne wyróżnienie tekstu pokazuje identyczność sekwencji (100%), a szare – wysoki stopień podobieństwa. **B** – 3 loga (Metody 4.24.) z wyznaczonymi zakonserwowanymi nukleotydami w analizowanym regionie oznaczone czarną ramką w konsensusie (A). Czerwoną ramką oznaczyłam proponowany motyw wiązania białka RacR. **C** – Proponowany motyw wiązania białka RacR do sekwencji intergenowej oraz swojej kodującej.

Z uwagi na wyznaczenie czterech sekwencji oddziaływania z DNA, gdzie każda pojedyncza sekwencja może być miejscem wiązania monomeru RacR, a palindromu – wiązaniem homodimeru RacR, postanowiliśmy zbadać status oligomeryczny białka RacR. W tym celu przeprowadzono sączenie molekularne na kolumnie Superdex 75 10/300 GL (Metody 4.14.). Najpierw wykonano doświadczenia niezbędne do wyznaczenia krzywej kalibracyjnej z wykorzystaniem zestawu do kalibracji filtrów żelowych dla niskiej masy cząsteczkowej (Ryc. 44. A). W wyniku rozdzielania próbki białka RacR uzyskano pojedynczy szczyt o objętości elucyjnej wynoszącej 8,21 ml o wyznaczonej masie 151,165 kDa (Ryc. 44. B). Biorąc pod uwagę sekwencję aminokwasową monomeru rekombinowanego białka RacR z obecnym znacznikiem 6his, to przewidywana masa cząsteczkowa wynosi dla monomeru 18,5 kDa. Otrzymany wynik pokazuje, że białko RacR w badanych warunkach przyjmuje formę jednorodnego oktameru w roztworze.



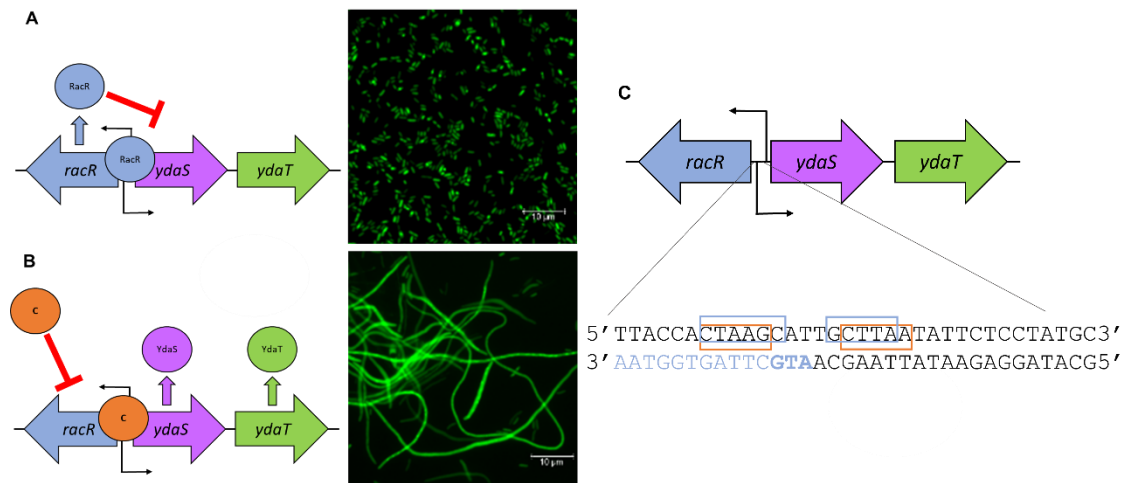
Ryc. 44. A – Sączenie molekularne białka RacR. Czerwoną strzałką oznaczyłam szczyt elucji dla białka RacR. Położenie wzorców masy cząsteczkowej pokazałam szarymi strzałkami. **B** – Krzywa kalibracyjna została wyznaczona poprzez rozdzielanie białek wzorcowych na kolumnie Superdex-75 10/300 GL (Citiva). Czerwonym kwadratem oznaczyłam masę białka RacR na krzywej.

6. Dyskusja

Reakcja krzyżowa (*ang. cross-talk*) pomiędzy dwoma bakteryjnymi czynnikami transkrypcyjnymi: regulatorem ekspresji systemu restrykcyjno-modyfikacyjnego oraz represorem defektywnego profaga

Powyższe badania wykazały ważną rolę czynników transkrypcyjnych (TF) pozyskanych na drodze horyzontalnego transferu genów w reakcjach modulowania ekspresji genów gospodarza. W dodatku, pokazały, że wnikający do komórek gospodarza DNA, nawet pochodzący z blisko spokrewnionego organizmu, jeśli koduje gen regulatora, to jest w stanie zakłócać sieć regulatorową gospodarza. Poprzez regulatorowy „cross-talk” zwykle rozumie się przypadkową reakcję krzyżową czynnika transkrypcyjnego z miejscem DNA, poza swoistym naturalnym celem (*ang. off-target site*) w odróżnieniu do specyficznej reakcji TF z naturalną sekwencją DNA (*ang. target site*). Zwykle TF, który wiąże DNA spoza swojego naturalnego celu określa się jako TF „niewierny”, czy o „rozluźnionej specyficzności” (*ang. promiscuous TF*) (Taylor et al., 2022). Oddziaływanie takiego TF z DNA może być ciągle reakcją specyficznego wiązania DNA, możliwe, że przypadkową i nie zawsze o znanej funkcji czy roli.

Na Rycinie 45A przedstawiłam warunki fizjologiczne dla badanego modelu operonu *racR – ydaST* w komórkach *E. coli*. Represor RacR wiąże się w regionie intergenowym blokując ekspresję genów *ydaS* oraz *ydaT*. Kiedy białko regulatorowe C zostaje wprowadzone do komórek, np. na drodze koniugacji, zaczyna wiązać się w sekwencji promotorowej genu *racR*. Przez to wiązanie zahamowana zostaje ekspresja genu *racR*. Z kolei niewystarczająca ilość represora RacR prowadzi do de-represji *ydaS* oraz *ydaT*, co w rezultacie prowadzi do powstawania filamentów komórkowych o nieznanym podłożu molekularnym (Ryc. 45. B). Moje wyniki wskazują na konkretne miejsce w genomie *E. coli*, w regionie profaga Rac, w którym regulon pod kontrolą białka RacR jest zakłócany i ten sygnał skutkuje zmianą fenotypu bakterii (Ryc. 45. C).



Ryc. 45. „Cross-talk” pomiędzy dwoma czynnikami transkrypcyjnymi: białkiem C pochodzącym z systemu R-M Csp23II oraz represorem defektywnego faga Rac. **A** – w warunkach fizjologicznych, gdy białko C nie jest obecne, represor RacR wiąże się w regionie intergenowym, co blokuje aktywność *ydaS* oraz *ydaT*. Ich ekspresja jest całkowicie wyciszone. **B** – Kiedy gen białka C zostaje wprowadzony do komórki na plazmidzie, jego ilość pozwala na związanie się do promotora własnego genu C, ale także zaczyna wiązać się w sekwencji genu *racR*. Wiązanie to prowadzi do zahamowania ekspresji genu *racR*. Brak represora RacR prowadzi do de-represji *ydaS* oraz *ydaT*. Aktywność YdaS oraz YdaT powoduje powstawanie filamentów prawdopodobnie przez przypuszczalną toksyczną aktywność białka YdaT. **C** – zmapowane miejsca zakłócania regulonu RacR w rejonie profaga Rac w genomie *E. coli*. Miejsce wiązania C oznaczyłam pomarańczową ramką, zaś wiązania RacR oznaczyłam niebieską ramką. Sekwencja kodująca gen *racR* jest na niebiesko, z pogrubionym startem transkrypcji.

E. coli K-12 MG1655 ma jeden z najlepiej zbadanych genomów (Hu et al., 2009), ale jeszcze nie wszystkie geny zostały scharakteryzowane (Gao et al., 2021). TF odgrywają najważniejszą rolę w regulacji ekspresji genów i funkcjonowaniu sieci genetycznych, mimo tego, również jeszcze nie wszystkie zostały scharakteryzowane (Gao et al., 2018). Czynniki transkrypcyjne *E. coli* K-12 zostały podzielone na 4 grupy na podstawie liczby docelowych miejsc regulacyjnych, które tworzą strukturę przypominającą sieć pajęczą. Regulacja tej sieci jest najważniejsza dla podstawowych procesów życiowych bakterii. Czynniki transkrypcyjne o zasięgu globalnym będą tworzyły wokół siebie gęstą sieć. Posiadają one wiele miejsc docelowych, w przeciwieństwie do czynników transkrypcyjnych o zasięgu lokalnym (Shimada et al., 2021; Yan et al., 2010). Kolejna grupa składa się z czynników transkrypcyjnych związanych z nukleoidami (NAP). Są one zaangażowane w organizację i utrzymanie struktury nukleoidu, co silnie wpływa na ekspresję genów (Kroner et al., 2019). Czynniki transkrypcyjne także odgrywają ważną rolę w adaptacji do warunków środowiska (van Hijum et al., 2009). W sumie kontrolują one ponad tysiąc genów, jednak trudno jest

przeanalizować efekt pojedynczego TF podczas badań laboratoryjnych, ze względu na cały repertuar współistniejących regulatorów (Ishihama, 2010). Ponadto każdy czynnik transkrypcyjny jest tak dostrojony, aby działać mniej specyficznym i wpływać na duże grupy genów (regulony) lub być wysoce specyficznym i działać w obrębie regulacji pojedynczego genu. Co ciekawe, większość TF rozpoznaje drugorzędowe miejsca docelowe, co powoduje większą złożoność systemów regulacyjnych i przyczynia się do powstawania reakcji krzyżowych typu „cross-talk” (Wunderlich & Mirny, 2009).

Zjawisko takie może wprowadzać zakłócenia w normalnie funkcjonującej komórce. Może mieć pozytywne lub negatywne implikacje dla komórki (Rowland et al., 2017). Jeżeli fragment DNA, który niesie TF, został wprowadzony poprzez horyzontalny transfer genów (HTG) do nowego gospodarza sprawa jest jeszcze bardziej skomplikowana. Wprowadzenie modułu genetycznego w tym wypadku na plazmidzie wymusza pełnienie swoich funkcji jednocześnie nie zakłócając działania istniejącej już sieci genetycznej gospodarza (Davids & Zhang, 2008).

Moja praca pokazuje przykład wzajemnego oddziaływania regulacyjnego pomiędzy czynnikiem transkrypcyjnym (białko C), które pochodzi z systemu R-M (*Citrobacter* sp.), a represorem RacR kontrolującym geny profaga Rac (*Escherichia coli*). Oba gatunki bakterii są blisko spokrewnione. Należą do rodziny *Enterobacteriaceae*. Odkrycie, że białko C wiąże się z regionem promotorowym układu R-M, jak i z miejscem *racR* spowodowało, że zajęłam się dalszą analizą tego zjawiska na poziomie molekularnym.

Operony niosące białko C w procesie HTG mogą zabijać lub obniżać żywotność nowych gospodarzy, którzy posiadają region Rac w swoim genomie. Badania pokazują, że tzw. niezgodność genów lub zakłócenie w funkcjonowaniu regulatorowych sieci zależności może stanowić barierę dla HTG (Thomas & Nielsen, 2005). Podobna niezgodność transkrypcyjna spowodowana przez tzw. „cross-talk” została opisana na przykładzie post-zygotycznej izolacji między gatunkami. Jest to rodzaj izolacji rozrodczej, która nie dopuszcza przepływu genów, gdy dojdzie już do zapłodnienia między blisko spokrewnionymi gatunkami (Tulchinsky et al., 2014). Jednak tamto zakłócenie ma na celu ograniczenie hybrydyzacji międzygatunkowej pomiędzy spokrewnionymi grupami organizmów. Ma on służyć jako swoistego rodzaju bariera gatunkowa. Reakcja krzyżowa (ang. *cross-talk*) została również opisana na mniejszą skalę – w celu rywalizacji samolubnych elementów genetycznych (ang. „*selfish genetic*

elements”). „Samolubny” cel to promowanie przetrwania ponad wszelkie inne korzyści (Kobayashi, 2001). Aby odpowiedzieć sobie na pytanie, czy jest możliwe współistnienie białka C (lub jego homologu) z genami *racR/ydaS/ydaT* została przeprowadzona analiza genomowa na grupie 495 zsekwencjonowanych genomów *E. coli*. Ogólnie zidentyfikowano homologu białka C tylko w 23 genomach ze wszystkich testowanych. Wśród tych wyszukanych 23, aż 16 z nich nie zawierało genów *racR* ani *ydaS*. Natomiast z tych 16 tylko jeden z nich posiadający homolog białka C posiadał również geny *racR/ydaS/ydaT*. Współwystępowanie wydaje się rzadkie (Negri et al., 2019). Może to być dowód na to, że te czynniki transkrypcyjne podczas reakcji krzyżowej (ang. *cross-talk*) są niezgodne i mogą powodować śmierć gospodarza, dlatego ich ko-egzystencja jest rzadka.

Docelowe miejsca wiązania DNA białek C to tak zwane C-boxy. Sekwencja konsensusowa dla dimeru C.Csp231I składa się z dwóch odwróconych powtórzeń (5'-CTAAG-n5-CTTAG-3') (Sorokin et al., 2009). Szukaliśmy takich miejsc w genomie *E. coli* i zostały znalezione, ale sekwencja rozdzielająca (n) pomiędzy powtórzeniami nie była równa 5 (Ryc. 46.).

Lokalizacja chromosom <i>E. coli</i>		Wyszukane sekwencje
1077641	1077654	CTAAGGGACTTAG
1144545	1144579	CTAAGTAATCACGCATCTGCGTGATGAAGCTTAG
2079989	2080010	CTAAGAGGACAGTCCTCCTTAG
3597007	3597021	CTAAGGGCGCTTAG
3617341	3617362	CTAAGAGTAGTGTGCTCCTTAG
3783082	3783094	CTAAGTACTTAG
3972924	3972956	CTAAGGCTATCCCTTTTGGCTGATAGCCTTAG
4235446	4235464	CTAAGGTTTTATCCTTAG

Ryc. 46. Potencjalne miejsca wiązania się białka C uwzględniając sekwencję konsensusową. Podkreślenie – sekwencja rozdzielająca powtórzenia.

Analizy transkryptomocne wykazały, że gen *racR* może odgrywać kluczową rolę regulatorową w ekspresji genu profaga Rac (Negri et al., 2019). Zauważyliśmy, że gen *racR* był jedynym genem, którego ekspresja się obniżyła, podczas ekspresji białka C, w przeciwieństwie do kilku genów sąsiadujących, których ekspresja została znacznie podwyższona (nawet 70 razy). Sekwencja aminokwasowa RacR (158 aa) nie jest podobna do białka C (98 aa) (Ryc. 47.) mimo, że oba białka oddziałują z zachodzącą na siebie sekwencją DNA (Ryc. 45 C). Oba białka to czynniki transkrypcyjne pochodzenia lambdoidalnego z motywem HTH do interakcji z DNA (Bindal et al., 2017).

C	MLIRRLKDARLRAGISQEKLGVLGIDEASASAR--MNQYEKGKHAPDFEMANRLAKVLK	58
RacR	-----MLSGKDLGRAIEQAINKKIASGSVKSKE-----VARHFK	35
	:*:* * . * : : : . * . * :*:* *	
C	I--PVSYLYPE-----D-----DLAQIIL	76
RacR	VQPPSIYDWIKKGSISKDKLPWLRFSDVVGPEHWGLNEYPIPTPTNSDTKSELLDINN	95
	: * * : : *	:* :*
C	TWNELEQERKRINFYIRKKAK-----	98
RacR	LYQAASDEIRAIIVAFLLSGNATEPDWVDHVRAYIAAMEMKVGKYLKALESERKSQNITK	155
	:: .:: * : * : :*.	
C	---	98
RacR	TGT	158

Ryc. 47. Zestawienie sekwencji aminokwasowej białka regulatorowego C i represora RacR.

Przeprowadzone przeze mnie testy *in vivo* wykazały wpływ białka C na ekspresję *racR*, a dokładniej wyciszenie jego ekspresji. Sugeruje się, że tzw. „cross-talk” pomiędzy czynnikiem transkrypcyjnym jest bardziej prawdopodobny w przypadku TF rozpoznających krótsze sekwencje DNA lub częściowo zdegenerowane (możliwe odstępstwo – mis-match od sekwencji consensus). W takim przypadku istnieje duże prawdopodobieństwo, że w genomie jest obecnych bardzo wiele miejsc, jednak o mniej specyficznym wiązaniu (Stewart & Plotkin, 2013). Ponadto docelowe miejsca wiązania są również narażone na stałą presję genetyczną, aby dostosować wiązanie TF poprzez mutacje punktowe, co napędza ewolucję sieci regulacyjnej (Payne & Wagner, 2014). Niemniej jednak spodziewaliśmy się znaleźć wiązanie się białka C do sekwencji podobnej do jej sekwencji konsensus. Znaleźliśmy jeszcze kilka sekwencji 5'-CTAAG-3' dla pojedynczego motywu rozpoznawanego przez białko C w obrębie sekwencji intergenowej pomiędzy *ydaS* i *racR*. Jednak tylko jedno miejsce odpowiadało sekwencji konsensus, blisko miejsca translacji genu *racR* (5'-CTAAG-n5-CTTAA-3'). Tylko jeden nukleotyd, który podkreśliłam jest różny od sekwencji konsensus. W swojej pracy przy użyciu metod EMSA i śladu DNA (ang. *footprinting*) potwierdziłam miejsce wiązania białka C z regionem faga *rac* (Ryc. 30,31,38,39). Hamowaniu ekspresji represora RacR może sprzyjać tendencja białek C do silnego zginania DNA (Mruk & Blumenthal, 2009). Ciekawe, że profil hamowania transkrypcji genu *racR*, w zależności od indukcji arabinozy dla białka C i RacR jest porównywalny, co potwierdziłam w teście reporterowym *in vivo* (Ryc. 44.). To nie musi oznaczać, że stężenia komórkowe białek są podobne przy tej samej ilości induktora arabinozy. Przewidujemy, że dimer białka C, a

nie monomer będzie oddziaływać z wyznaczoną sekwencją w teście protekcji DNazyI (Ryc. 38, 39). Co ciekawe, wydaje się, że miejsce wiązania białka C i RacR częściowo nakładają się na siebie: 5'-ACTAAGCATTGCTTAATATTCTC-3'. Podkreślone zostało miejsce sekwencji konsensusu dla białka C, a pogrubione dla RacR (Ryc. 45.). Ponadto nasze inne badania pokazują, że gdy oba białka są obecne razem, inicjacja transkrypcji *in vitro* z promotora *ydaS* nie jest już w pełni hamowana przez represor RacR. Promotor *ydaS* (wspólny też dla *ydaT*) jest bardzo silnym promotorem, co wynika z analizy sekwencji oraz wyników transkrypcji *in vitro* (Wisniewska et al., 2022). Jednak *in vivo* jest ledwie wykrywalny (Ryc. 29.). Wyniki te mogą odzwierciedlać wielowarstwowe interakcje między białkami C i RacR oraz regionem IGR pomiędzy genami *ydaS* i *racR*.

Istnieją publikacje, które bazują na modelach teoretycznych, mających na celu określenie możliwych scenariuszy interakcji reakcji krzyżowych. Ponadto mają sprawdzić, czy niewłaściwa regulacja genów może wpływać na adaptacyjną ekspresję genów (Friedlander et al., 2016; Wagner, 2021). W rzeczywistości dynamika interakcji pomiędzy TF tworzy równowagę pomiędzy utrzymaniem specyficzności wiązania, a umożliwieniem tworzenia nowych interakcji, aby umożliwić adaptację sieci. Jeśli nowo powstałe połączenia (ang. *rewiring*) TF są korzystne to ułatwią ewolucję sieci (Taylor et al., 2022). Czasami jest tak, jak w naszym przypadku, że wyzwala toksyczną aktywność genów, która jest wyciszona w normalnych warunkach (Negri et al., 2019). Istnieje również inny udokumentowany przykład białek regulatorowych C występujących w systemie R-M Typu I pochodzącego z bakterii *Mycoplasma*. Białko to również wykazuje zdolność do interakcji z celami poza swoją naturalną sekwencją i również w tym przypadku interakcje te są szkodliwe dla komórki bakteryjnej. Odkryto, że białko C wiąże przewidywane powtórzenia motywu 5'-GTGTTA-N₅-GTGTTA-3'. Oprócz własnego promotora operonu HsdC (naturalny cel) może wiązać się z promotorami genu białka opiekuńczego *clpB* i genem tRNA-Met (gen *trnM*) (Fisunov et al., 2017). Badania pokazują, że dodanie nowego połączenia regulatorowego do istniejącej sieci może powodować duże zakłócenia (Baumstark et al., 2015). Horyzontalny transfer genów z udziałem TF jest zjawiskiem powszechnym, dlatego istnieje presja genetyczna na sieć regulacyjną, aby zapobiec powstawaniu zaburzeń (Rowland et al., 2017).

Badane przeze mnie białko C blokowało transkrypcję genu *racR* na etapie inicjacji transkrypcji. Jednak taką samą funkcję można osiągnąć poprzez blokowanie transkrypcji

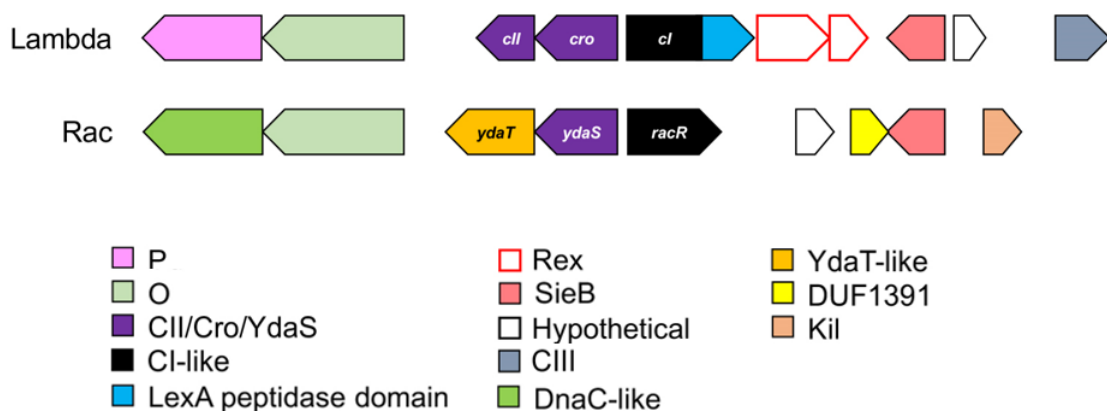
podczas etapu elongacji transkrypcji (Klimuk et al., 2018). W systemie R-M Kpn2I, które również zawiera białko regulatorowe C zmapowano dwa miejsca wiązania w DNA, ale w obrębie swojej sekwencji kodującej genu C, blisko kodonu inicjatorowego ATG, co prawdopodobnie tworzy silną blokadę dla wydłużającego się kompleksu polimerazy RNA.

Większość zsekwencjonowanych genomów *E. coli* ma przynajmniej część sekwencji rac (Krishnamurthi et al., 2017). Region rac zawiera 29 genów, chociaż niektóre z nich nie zostały scharakteryzowane, a ich powiązania genetyczne nie zostały określone (Ghatak et al., 2019). Profag Rac był pierwszym defektywnym profagiem lambdoidalnym odkrytym w genomie *E. coli* K-12 (Krishnamurthi et al., 2017). Jego rozmiar i zawartość genów różnią się w zależności od szczepów *E. coli*. Ocenia się, że prawdopodobnie utracił 60% swojej pierwotnej zawartości DNA. Większość z zsekwencjonowanych genomów *E. coli* ma przynajmniej część genów rac (Casjens, 2003). Tylko kilka z nich jest wysoce konserwowanych w tym gen *recE*, zaangażowany w alternatywny szlak rekombinacji homologicznej oraz *trkG* kodujący permeazę jonów potasu (Krishnamurthi et al., 2017). Gen represora RacR zawiera motyw HTH (helisa-skret-helisa) i jest spokrewniony z represorem lambda CI (15% identyczności według dopasowania Needlemana-Wunscha). Jego usunięcie jest nieprzeżyciowe (ang. *essential gene*) (Krishnamurthi et al., 2017). Kolekcja Keio mutantów delecyjnych pojedynczych genów *E. coli* nie zawiera szczepu $\Delta racR$ (Baba et al., 2006). Wydaje się, że represor RacR jest głównym regulatorem sieci ekspresyjnej Rac, która wraz z sąsiednimi genami *ydaS* i *ydaT* tworzy operon w genomie *E. coli*, uznawany za niezbędny (Krishnamurthi et al., 2017; Bindal et al., 2017; Sorek et al., 2007). Profag Rac jest wysoce konserwowany w genomie *E. coli* (Casjens, 2003). Badania pokazują, że profag ten i inne profagi pomagają komórkom *E. coli* wytrzymać stres osmotyczny, oksydacyjny i kwasowy tworząc biofilm (Wang et al., 2010).

Analogie rejonu *racR* – *ydaST* do regionu „decyzji” bakteriofaga λ

Wśród 29 genów regionu rac występują także *ydaS* oraz *ydaT*, które posiadają tylko domniemane adnotacje dotyczące funkcji (Jobling, 2018). Zasugerowano, że oba działają w podobny sposób jak represory Cro i CII, które koordynują ekspresję genów bakteriofaga lambda związaną z tzw. decyzją wejścia w cykl lityczny, czy lizogeniczny (Casjens, 2003). Jednak tylko białko YdaS zawiera typowy motyw HTH. Możliwe, że

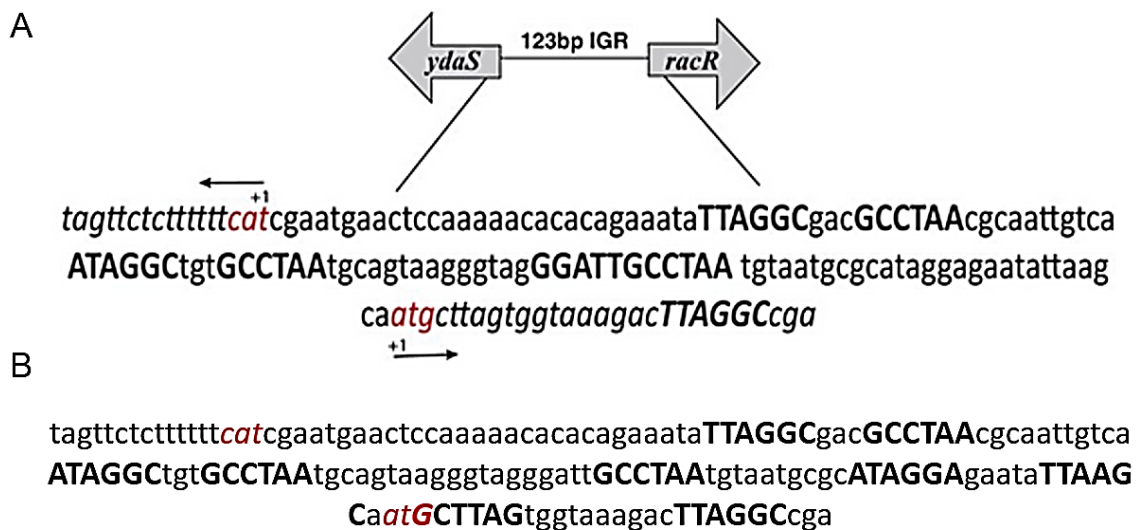
toksyczna aktywność operonu *ydaST* wynika z obecności ważnego przełącznika genetycznego (ang. *genetic switches*) dla *ydaS* (regulatora) w celu aktywacji genu *ydaT* poprzez wspólny promotor. Toksyczność komórkowa, objawiająca się rozległą filamentacją komórek może zatem pochodzić wyłącznie od białka YdaT. To pokazuje nieodzowną rolę represora RacR w sieci regulowania ekspresji genów profaga Rac. Mechanizm toksyczności YdaT nie jest zrozumiały, podobnie jak jego funkcja w sieci profaga Rac. W trakcie naszej pracy znaleźliśmy bardzo bliską analogię regionu IGR zlokalizowanego między genami *racR* i *ydaST* do tzw. regionu decyzyjnego liza/lizogenia pomiędzy genami CI i Cro bakteriofaga lambda (Ryc. 48.) (Ptashne, 2011). Zostało to również zauważone przez innych badaczy (Krishnamurthi et al., 2017; Jobling, 2018; Jurenas et al., 2021).



Ryc. 48. Schematyczne przedstawienie podobieństwa genów wchodzących w skład faga Rac do Lambdy. Opracowałam na podstawie (Jurenas et al., 2021).

Potwierdzamy, że RacR ma podobieństwa funkcjonalne do represora λ CI (jest także oktamerem w roztworze), a także YdaS przypominające białko λ Cro, które promuje przejście transkrypcji faga λ do cyklu litycznego. YdaS działa jako autorepresor, ale wydaje się, że RacR silniej hamuje ekspresję *ydaS*, pomimo że promotory *racR* wydają się być słabsze niż pojedynczy promotor *ydaS*. Wiedza o sile promotorów wynika z badań transkrypcji *in vitro* (Wisniewska et al., 2022). Skuteczne hamowanie *ydaS* zaobserwowaliśmy przy bardzo niskim stężeniu RacR zarówno w eksperymentach *in vivo*, jak i *in vitro* (Wisniewska et al., 2022 Ryc. 2 i 7). Fakt ten może wskazywać, że RacR ma większe powinowactwo do operatora 1 w obrębie sekwencji promotora *ydaS*. Jest to kolejne podobieństwo jak wiązanie Cro/CI do operatora (O_R) (Ptashne, 2011). Wykorzystując metodę śladu DNA (ang. *footprinting*) zidentyfikowałam dokładną

lokalizację miejsca wiązania się represora RacR, określoną jako trzy odwrócone powtórzenia (operatory), które nieznacznie różnią się między sobą sekwencją (Ryc. 43.), podobnie jak trzy operatory λ OR dla CI i Cro (Ptashne, 2011). Jedna z publikacji wyznacza wiązanie się RacR do sekwencji 5'-TTAGGC-n3-GCCTAA-3' oraz pokazuje, że występują wymienione sekwencje w trzech miejscach regionu IGR – w sumie 6 miejsc wiązania (Ryc. 49.). Wykonany został również test EMSA pokazujący wiązanie się białka RacR do IGR (Krishnamurthi et al., 2017). Zgadza się z tym, że wiązanie tam występuje natomiast opierając się na moich badaniach *in vitro* (Ryc. 43.) uważam, że znajdują się tam trzy miejsca wiązania (trzy operatory, każdy wiąże dimer RacR), a ponadto czwarty operator w sekwencji kodującej *racR* wydaje się najbardziej różny od sekwencji konsensus, ale tam również występuje wiązanie (Ryc. 43.).



Ryc. 49. Sekwencja DNA regionu intergenowego pomiędzy genami *racR* oraz *ydaS* (123 p.z.). **A** – sekwencja IGR z zaproponowanym miejscem wiązania RacR. Opracowałam na podstawie Krishnamurthi et al., 2017 **B** – sekwencja IGR z miejscem wiązania wyznaczonym na podstawie badań śladu DNA (ang. *footprinting*). Na czerwono oznaczyłam +1 *ydaS* oraz +1 *racR*. Duże litery oznaczają wyznaczone sekwencje wiązania RacR.

Moje wyniki wskazują na duże podobieństwo profaga Rac do λ , jednakże występują także różnice. λ Cro hamuje ekspresję represora CI. YdaS nie wydaje się osłabiać ekspresji *racR* (Ryc. 30.). Pomimo występowania podobieństw w organizacji promotora i operatora obu fagów, różnią się one regulacją. Ponadto miejsce operatora 3 (oznaczonego przeze mnie jako O₃ Ryc. 43. A,B) *racR* nakłada się na sekwencję kodującą genu *racR*.

Najnowsze badania nad profagiem lambdoidowym *E. coli* CP-933P, który jest podobny do profaga Rac, również sugerują analogię do λ (Jurenas et al., 2021). Wydaje się, że PaaR2 jest analogiem czynnika transkrypcyjnego RacR. Tam miejsca wiązania

PaaR2 są obecne w regionie międzygenowym pomiędzy genami *paar2* i *ydaS*, tworząc cztery odwrócone powtórzenia, które nie nakładają się na sekwencję kodującą. Przeciwnieństwem jest aktywność białka YdaT. Białko profaga CP-933P nie ma żadnych skutków toksycznych (Jurenas et al., 2021). YdaT pochodzące od profaga Rac prowadzi do filamentacji komórek, co nie wynika z indukcji odpowiedzi SOS (Ryc. 11.). Regulacja profaga Rac i funkcje obejmujących go genów, różnią się od innych systemów opisanych do tej pory oraz wymagają dalszych badań.

7. Podsumowanie

Powyższe badania wykazały ważną rolę czynników transkrypcyjnych (TF) pozyskiwanych na drodze horyzontalnego transferu genów w reakcjach modulowania ekspresji genów gospodarza na poziomie globalnym. W toku naszych badań zauważyliśmy, że wprowadzenie na plazmidzie genów systemu restrykcyjno-modyfikacyjnego (R-M) Typu II - Csp231I, skutkuje ciekawym zjawiskiem tworzenia się drastycznie wydłużonych komórek *E. coli*.

- Ustaliłam, że za zjawisko to odpowiedzialne jest białko regulatorowe C, o aktywności czynnika transkrypcyjnego, autorepresora, dedykowanego do regulacji ekspresji genów systemu R-M.
- Białko C wiąże się do sekwencji tzw. C-box w obrębie swojego promotora, ale także jak wykazałam jest zdolne do wiązania się do sekwencji genu *racR* w rejonie profaga Rac w genomie *E. coli* K-12 MG1655.
- Określiłam dodatkowe (poza naturalnym celem) miejsce wiązania regulatora C: 5'-CTAAGCATTGCTTAA-3' i wykazałam, że ta sekwencja jest też miejscem wiązania się białka RacR o aktywności autorepresora oraz jednocześnie represora sąsiednich genów: *ydaS* oraz *ydaT* o nieznannej funkcji.
- Wykazałam, że wiązanie białka C skutkuje zakłóceniem funkcjonowania regulonu RacR, tzn. prowadzi do znacznego spadku ekspresji genu *racR*, co z kolei wpływa na de-represję *ydaS* i *ydaT*.
- Aktywność YdaS i YdaT (przypuszczalnych czynników transkrypcyjnych) prowadzi do formowania się filamentów komórkowych, w procesie niezależnym od komórkowej odpowiedzi SOS, i związane jest prawdopodobnie z zaburzeniem podziałów komórkowych i segregacji komórek potomnych.
- Ustaliłam, że ekspresja regulatora C zmniejsza żywotność komórek *E. coli* MG1655, a nawet prowadzi do ich śmierci. Jednakże, w ciągu kilkudziesięciu generacji komórka spontanicznie jest w stanie zmienić fenotyp i zatrzymać proces filamentacji na drodze nieznanego mechanizmu.
- Nie jest jasne, czy zjawisko reakcji krzyżowej pomiędzy czynnikami transkrypcyjnymi: białka C i represora RacR są przypadkowe, ale otrzymane wyniki pokazują przykład zdarzenia horyzontalnego transferu genów, które może prowadzić do zmniejszenia żywotności komórek gospodarza, a nawet jego śmierci. Być może zjawiska te są częste, lecz niedostrzegane i przez to nie dość wnikliwie badane.

8. Literatura

- Alekshun, M. N., & Levy, S. B. (1997). Regulation of chromosomally mediated multiple antibiotic resistance: the *mar* regulon. *Antimicrobial Agents and Chemotherapy*, *41*(10), 2067–2075.
- Anderson, J. J., Quay, S. C., & Oxender, D. L. (1976). Mapping of two loci affecting the regulation of branched-chain amino acid transport in *Escherichia coli* K-12. *Journal of Bacteriology*, *126*(1), 80–90.
- Atlung, T., & Ingmer, H. (1997). H-NS: a modulator of environmentally regulated gene expression. *Molecular Microbiology*, *24*(1), 7–17.
- Azam, T. A., Iwata, A., Nishimura, A., Ueda, S., & Ishihama, A. (1999). Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *Journal of Bacteriology*, *181*(20), 6361–6370.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, *2*, 2006.0008.
- Babu, M. M., & Teichmann, S. A. (2003). Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucleic Acids Research*, *31*(4), 1234–1244.
- Bachmann, B. J. (1972). Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriological Reviews*, *36*(4), 525–557.
- Bae, S. H., Liu, D., Lim, H. M., Lee, Y., & Choi, B. S. (2008). Structure of the nucleoid-associated protein Cnu reveals common binding sites for H-NS in Cnu and Hha. *Biochemistry*, *47*(7), 1993–2001.
- Baumstark, R., Hänzelmann, S., Tsuru, S., Schaerli, Y., Francesconi, M., Mancuso, F. M., Castelo, R., & Isalan, M. (2015). The propagation of perturbations in rewired bacterial gene networks. *Nature Communications*, *6*.
- Bessaiah, H., Anamalé, C., Sung, J., & Dozois, C. M. (2021). What Flips the Switch? Signals and Stress Regulating Extraintestinal Pathogenic *Escherichia coli* Type 1 Fimbriae (Pili). *Microorganisms 2022, Vol. 10, Page 5*, *10*(1), 5.
- Bindal, G., Krishnamurthi, R., Seshasayee, A. S. N., & Rath, D. (2017). CRISPR-Cas-Mediated Gene Silencing Reveals RacR To Be a Negative Regulator of YdaS and YdaT Toxins in *Escherichia coli* K-12. *MSphere*, *2*(6).
- Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heynker, H. L., Boyer, H. W., Crosa, J. H., & Falkow, S. (1992). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. 1977. *Biotechnology (Reading, Mass.)*, *24*(2), 153–171. <https://pubmed.ncbi.nlm.nih.gov/344137/>

- Boyle, D. S., Khattar, M. M., Addinall, S. G., Lutkenhaus, J., & Donachie, W. D. (1997). *ftsW* is an essential cell-division gene in *Escherichia coli*. *Molecular Microbiology*, *24*(6), 1263–1273.
- Browning, D. F., & Busby, S. J. W. (2004). The regulation of bacterial transcription initiation. *Nature Reviews Microbiology* *2004* *2*:1, *2*(1), 57–65.
- Burgess, R. R. (2001). Sigma Factors. *Encyclopedia of Genetics*, 1831–1834.
- Calvo, J. M., & Matthews, R. G. (1994). The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiological Reviews*, *58*(3), 466–490.
- Capra, E. J., Perchuk, B. S., Skerker, J. M., & Laub, M. T. (2012). Adaptive mutations that prevent cross-talk enable the expansion of paralogous signaling protein families. *Cell*, *150*(1), 222.
- Casadaban, M. J., & Cohen, S. N. (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *Journal of Molecular Biology*, *138*(2), 179–207.
- Casjens, S. (2003). Prophages and bacterial genomics: what have we learned so far? *Molecular Microbiology*, *49*(2), 277–300.
- Chen, S., Rosner, M. H., & Calvo, J. M. (2001). Leucine-regulated self-association of leucine-responsive regulatory protein (Lrp) from *Escherichia coli*. *Journal of Molecular Biology*, *312*(4), 625–635.
- Cho, B. K., Barrett, C. L., Knight, E. M., Park, Y. S., & Palsson, B. (2008). Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19462–19467.
- Cho, B. K., Knight, E. M., Barrett, C. L., & Palsson, B. (2008). Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Research*, *18*(6), 900.
- Chubiz, L. M., Glekas, G. D., & Rao, C. v. (2012). Transcriptional Cross Talk within the mar-sox-rob Regulon in *Escherichia coli* Is Limited to the *rob* and *marRAB* Operons. *Journal of Bacteriology*, *194*(18), 4867.
- Cramer, P. (2002). Multisubunit RNA polymerases. *Current Opinion in Structural Biology*, *12*(1), 89–97.
- Craven, S. J., Condon, S. G. F., Vázquez, G. D., Cui, Q., & Senes, A. (2022). The coiled-coil domain of *Escherichia coli* FtsLB is a structurally detuned element critical for modulating its activation in bacterial cell division. *Journal of Biological Chemistry*, *298*(1), 101460.
- Davids, W., & Zhang, Z. (2008). The impact of horizontal gene transfer in shaping operons and protein interaction networks--direct evidence of preferential attachment. *BMC Evolutionary Biology*, *8*(1).

- Diederich, L., Roth, A., & Messer, W. (1994). A versatile plasmid vector system for the regulated expression of genes in *Escherichia coli*. *BioTechniques*, *16*(5), 916-918+923. <https://pubmed.ncbi.nlm.nih.gov/8068348/>
- Dorman, C. J. (2004). H-NS: a universal regulator for a dynamic genome. *Nature Reviews. Microbiology*, *2*(5), 391–400.
- Feng, Y., & Cronan, J. E. (2012). Crosstalk of *Escherichia coli* FadR with global regulators in expression of fatty acid transport genes. *PLoS ONE*, *7*(9), 46275.
- Fisunov, G. Y., Evsyutina, D. v., Manuvera, V. A., & Govorun, V. M. (2017). Binding site of restriction-modification system controller protein in Mollicutes. *BMC Microbiology*, *17*(1), 1–7.
- Foo, Y. H., Spahn, C., Zhang, H., Heilemann, M., & Kenney, L. J. (2015). Single cell super-resolution imaging of *E. coli* OmpR during environmental stress. *Integrative Biology : Quantitative Biosciences from Nano to Macro*, *7*(10), 1297–1308.
- Friedlander, T., Prizak, R., Guet, C. C., Barton, N. H., & Tkacik, G. (2016). Intrinsic limits to gene regulation by global crosstalk. *Nature Communications*, *7*.
- Gama-Castro, S., Jiménez-Jacinto, V., Peralta-Gil, M., Santos-Zavaleta, A., Peñaloza-Spinola, M. I., Contreras-Moreira, B., Segura-Salazar, J., Muñiz-Rascado, L., Martínez-Flores, I., Salgado, H., Bonavides-Martínez, C., Abreu-Goodger, C., Rodríguez-Penagos, C., Miranda-Ríos, J., Morett, E., Merino, E., Huerta, A. M., Treviño-Quintanilla, L., & Collado-Vides, J. (2008). RegulonDB (version 6.0): gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nucleic Acids Research*, *36*(Database issue), D120.
- Gao, Y., Lim, H. G., Verkler, H., Szubin, R., Quach, D., Rodionova, I., Chen, K., Yurkovich, J. T., Cho, B. K., & Palsson, B. O. (2021). Unraveling the functions of uncharacterized transcription factors in *Escherichia coli* using ChIP-exo. *Nucleic Acids Research*, *49*(17), 9696–9710.
- Gao, Y., Yurkovich, J. T., Seo, S. W., Kabimoldayev, I., Dräger, A., Chen, K., Sastry, A. v., Fang, X., Mih, N., Yang, L., Eichner, J., Cho, B. K., Kim, D., & Palsson, B. O. (2018). Systematic discovery of uncharacterized transcription factors in *Escherichia coli* K-12 MG1655. *Nucleic Acids Research*, *46*(20), 10682–10696.
- Ghatak, S., King, Z. A., Sastry, A., & Palsson, B. O. (2019). The y-ome defines the 35% of *Escherichia coli* genes that lack experimental evidence of function. *Nucleic Acids Research*, *47*(5), 2446–2454.
- Grainger, D. C., Hurd, D., Harrison, M., Holdstock, J., & Busby, S. J. W. (2005). Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(49), 17693–17698.

- Grimbs, A., Klosik, D. F., Bornholdt, S., & Hütt, M. T. (2019). A system-wide network reconstruction of gene regulation and metabolism in *Escherichia coli*. *PLOS Computational Biology*, *15*(5), e1006962.
- Gruber, T. M., & Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology*, *57*, 441–466.
- Guo, Y., Quiroga, C., Chen, Q., McAnulty, M. J., Benedik, M. J., Wood, T. K., & Wang, X. (2014). RalR (a DNase) and RalA (a small RNA) form a type I toxin–antitoxin system in *Escherichia coli*. *Nucleic Acids Research*, *42*(10), 6448–6462.
- Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, *177*(14), 4121–4130.
- Handa, N., Ichige, A., Kusano, K., & Kobayashi, I. (2000). Cellular responses to postsegregational killing by restriction- modification genes. *Journal of Bacteriology*, *182*(8), 2218–2229.
- Harari, O., Park, S. Y., Huang, H., Groisman, E. A., & Zwir, I. (2010). Defining the Plasticity of Transcription Factor Binding Sites by Deconstructing DNA Consensus Sequences: The PhoP-Binding Sites among Gamma/Enterobacteria. *PLOS Computational Biology*, *6*(7), e1000862.
- Holden, N. J., & Gally, D. L. (2004). Switches, cross-talk and memory in *Escherichia coli* adherence. *Journal of Medical Microbiology*, *53*(7), 585–593.
- Holland, A. M., & Rather, P. N. (2008). Evidence for extracellular control of RpoS proteolysis in *Escherichia coli*. *FEMS Microbiology Letters*, *286*(1), 50–59.
- Hu, P., Janga, S. C., Babu, M., Díaz-Mejía, J. J., Butland, G., Yang, W., Pogoutse, O., Guo, X., Phanse, S., Wong, P., Chandran, S., Christopoulos, C., Nazarians-Armavil, A., Nasser, N. K., Musso, G., Ali, M., Nazemof, N., Eroukova, V., Golshani, A., ... Emili, A. (2009). Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biology*, *7*(4), 0929–0947.
- Huang, H., Shao, X., Xie, Y., Wang, T., Zhang, Y., Wang, X., & Deng, X. (2019). An integrated genomic regulatory network of virulence-related transcriptional factors in *Pseudomonas aeruginosa*. *Nature Communications* 2019 10:1, *10*(1), 1–13.
- Hugonnet, J. E., Mengin-Lecreulx, D., Monton, A., den Blaauwen, T., Carbonnelle, E., Veckerlé, C., Yves, V. B., van Nieuwenhze, M., Bouchier, C., Tu, K., Rice, L. B., & Arthur, M. (2016). Factors essential for L,D-transpeptidase-mediated peptidoglycan cross-linking and β -lactam resistance in *Escherichia coli*. *ELife*, *5*(OCTOBER2016).
- Igler, C., Lagator, M., Tkačik, G., Bollback, J. P., & Guet, C. C. (2018). Evolutionary potential of transcription factors for gene regulatory rewiring. *Nature Ecology & Evolution*, *2*(10), 1633–1643.
- Ishihama, A. (2009). The Nucleoid: an Overview. *EcoSal Plus*, *3*(2).

- Ishihama, A. (2010). Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiology Reviews*, 34(5), 628–645.
- Jensen, K. F. (1993). The *Escherichia coli* K-12 'wild types' W3110 and MG1655 have an *rph* frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels. *Journal of Bacteriology*, 175(11), 3401–3407.
- Jishage, M., & Ishihama, A. (1998). A stationary phase protein in *Escherichia coli* with binding activity to the major σ subunit of RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 95(9), 4953.
- Jishage, M., Kvint, K., Shingler, V., & Nyström, T. (2002). Regulation of σ factor competition by the alarmone ppGpp. *Genes & Development*, 16(10), 1260.
- Jobling, M. G. (2018). Ectopic Expression of the *ydaS* and *ydaT* Genes of the Cryptic Prophage Rac of *Escherichia coli* K-12 May Be Toxic but Do They Really Encode Toxins?: a Case for Using Genetic Context To Understand Function. *MSphere*, 3(2).
- Jurenas, D., Fraikin, N., Goormaghtigh, F., de Bruyn, P., Vandervelde, A., Zedek, S., Jové, T., Charlier, D., Loris, R., & van Melderen, L. (2021). Bistable Expression of a Toxin-Antitoxin System Located in a Cryptic Prophage of *Escherichia coli* O157:H7. *MBio*, 12(6).
- Kahramanoglou, C., Seshasayee, A. S. N., Prieto, A. I., Ibberson, D., Schmidt, S., Zimmermann, J., Benes, V., Fraser, G. M., & Luscombe, N. M. (2011). Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Research*, 39(6), 2073.
- Klimuk, E., Bogdanova, E., Nagornykh, M., Rodic, A., Djordjevic, M., Medvedeva, S., Pavlova, O., & Severinov, K. (2018). Controller protein of restriction-modification system Kpn2I affects transcription of its gene by acting as a transcription elongation roadblock. *Nucleic Acids Research*, 46(20), 10810–10826.
- Kobayashi, I. (2001). Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Research*, 29(18), 3742–3756.
- Kotelnikova, E. A., Makeev, V. J., & Gelfand, M. S. (2005). Evolution of transcription factor DNA binding sites. *Gene*, 347(2), 255–263.
- Krishnamurthi, R., Ghosh, S., Khedkar, S., & Seshasayee, A. S. N. (2017). Repression of YdaS Toxin Is Mediated by Transcriptional Repressor RacR in the Cryptic rac Prophage of *Escherichia coli* K-12. *MSphere*, 2(6).
- Kroner, G. M., Wolfe, M. B., & Freddolino, P. L. (2019). *Escherichia coli* Lrp Regulates One-Third of the Genome via Direct, Cooperative, and Indirect Routes. *Journal of Bacteriology*, 201(3).
- Kurafeiski, J. D., Pinto, P., & Bornberg-Bauer, E. (2019). Evolutionary potential of cis-regulatory mutations to cause rapid changes in transcription factor binding. *Genome Biology and Evolution*, 11(2), 406–414.

- Lamrabet, O., Plumbridge, J., Martin, M., Lenski, R. E., Schneider, D., & Hindre, T. (2019). Plasticity of promoter-core sequences allows bacteria to compensate for the loss of a key global regulatory gene. *Molecular Biology and Evolution*, *36*(6), 1121–1133.
- Lin, R., D'Ari, R., & Newman, E. B. (1992). Lambda placMu insertions in genes of the leucine regulon: extension of the regulon to genes not regulated by leucine. *Journal of Bacteriology*, *174*(6), 1948–1955.
- Lin, S. H., & Lee, J. C. (2003). Determinants of DNA bending in the DNA-cyclic AMP receptor protein complexes in *Escherichia coli*. *Biochemistry*, *42*(17), 4809–4818.
- McAdams, H. H., Srinivasan, B., & Arkin, A. P. (2004). The evolution of genetic regulatory systems in bacteria. *Nature Reviews. Genetics*, *5*(3), 169–178.
- McFarland, K. A., Lucchini, S., Hinton, J. C. D., & Dorman, C. J. (2008). The leucine-responsive regulatory protein, Lrp, activates transcription of the *fim* operon in *Salmonella enterica* serovar typhimurium via the *fimZ* regulatory gene. *Journal of Bacteriology*, *190*(2), 602–612.
- Mercer, K. L. N., & Weiss, D. S. (2002). The *Escherichia coli* Cell Division Protein FtsW Is Required To Recruit Its Cognate Transpeptidase, FtsI (PBP3), to the Division Site. *Journal of Bacteriology*, *184*(4), 904.
- Monteiro, L. M. O., Sanches-Medeiros, A., Westmann, C. A., & Silva-Rocha, R. (2019). Modulating Fis and IHF binding specificity, crosstalk and regulatory logic through the engineering of complex promoters. *BioRxiv*, 614396.
- Mruk, I., & Blumenthal, R. M. (2009). Tuning the relative affinities for activating and repressing operators of a temporally regulated restriction-modification system. *Nucleic Acids Research*, *37*(3), 983–998.
- Mruk, I., & Kaczorowski, T. (2007). A rapid and efficient method for cloning genes of type II restriction-modification systems by use of a killer plasmid. *Applied and Environmental Microbiology*, *73*(13), 4286–4293.
- Mruk, I., & Kobayashi, I. (2014). To be or not to be: regulation of restriction-modification systems and other toxin-antitoxin systems. *Nucleic Acids Research*, *42*(1), 70–86.
- Negri, A., Jąkański, M., Szczuka, A., Prysycz, L. P., & Mruk, I. (2019). Transcriptome analyses of cells carrying the Type II Csp231I restriction–modification system reveal cross-talk between two unrelated transcription factors: C protein and the Rac prophage repressor. *Nucleic Acids Research*, *47*(18), 9542–9556.
- Payne, J. L., & Wagner, A. (2014). The robustness and evolvability of transcription factor binding sites. *Science (New York, N.Y.)*, *343*(6173), 875–877.
- Perez-Rueda, E., Hernandez-Guerrero, R., Martinez-Nuñez, M. A., Armenta-Medina, D., Sanchez, I., & Ibarra, J. A. (2018). Abundance, diversity and domain architecture variability in prokaryotic DNA-binding transcription factors. *PLOS ONE*, *13*(4), e0195332.

- Pleška, M., Qian, L., Okura, R., Bergmiller, T., Wakamoto, Y., Kussell, E., & Guet, C. C. (2016). Bacterial autoimmunity due to a restriction-modification system. *Current Biology : CB*, *26*(3), 404–409.
- Ptashne, M. (2011). Principles of a switch. *Nature Chemical Biology*, *7*(8), 484–487.
- Reyes-Fernández, E. Z., & Schuldiner, S. (2020). Acidification of cytoplasm in *Escherichia coli* provides a strategy to cope with stress and facilitates development of antibiotic resistance. *Scientific Reports*, *10*(1).
- Rezulak, M., Borsuk, I., & Mruk, I. (2016). Natural C-independent expression of restriction endonuclease in a C protein-associated restriction-modification system. *Nucleic Acids Research*, *44*(6), 2646–2660.
- Roberts, R. J., Belfort, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S. K., Dryden, D. T. F., Dybvig, K., Firman, K., Gromova, E. S., Gumpert, R. I., Halford, S. E., Hattman, S., Heitman, J., Hornby, D. P., Janulaitis, A., Jeltsch, A., ... Xu, S. Y. (2003). A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Research*, *31*(7), 1805–1812.
- Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. (2010). REBASE--a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research*, *38*(Database issue).
- Roesch, P. L., & Blomfield, I. C. (1998). Leucine alters the interaction of the leucine-responsive regulatory protein (Lrp) with the *fim* switch to stimulate site-specific recombination in *Escherichia coli*. *Molecular Microbiology*, *27*(4), 751–761.
- Rosanova, A., Colliva, A., Osella, M., & Caselle, M. (2017). Modelling the evolution of transcription factor binding preferences in complex eukaryotes. *Scientific Reports* *2017* *7*:1, *7*(1), 1–12.
- Rowland, M. A., Abdelzaher, A., Ghosh, P., & Mayo, M. L. (2017). Crosstalk and the dynamical modularity of feed-forward loops in transcriptional regulatory networks. *Biophysical Journal*, *112*(8), 1539–1550.
- Schneider, R., Travers, A., & Muskhelishvili, G. (1997). FIS modulates growth phase-dependent topological transitions of DNA in *Escherichia coli*. *Molecular Microbiology*, *26*(3), 519–530.
- Seshasayee, A. S. N., Sivaraman, K., & Luscombe, N. M. (2011). An overview of prokaryotic transcription factors : a summary of function and occurrence in bacterial genomes. *Sub-Cellular Biochemistry*, *52*, 7–23.
- Shen-Orr, S. S., Milo, R., Mangan, S., & Alon, U. (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics*, *31*(1), 64–68.
- Shevtsov, M. B., Streeter, S. D., Thresh, S. J., Swiderska, A., McGeehan, J. E., & Kneale, G. G. (2015). Structural analysis of DNA binding by C.Csp231I, a member of a novel class of R-M

controller proteins regulating gene expression. *Acta Crystallographica Section D: Biological Crystallography*, 71(Pt 2), 398–407.

Shimada, T., Ogasawara, H., Kobayashi, I., Kobayashi, N., & Ishihama, A. (2021). Single-target regulators constitute the minority group of transcription factors in *Escherichia coli* K-12. *Frontiers in Microbiology*, 12, 1646.

Shin, M., Song, M., Joon, H. R., Hong, Y., Kim, Y. J., Seok, Y. J., Ha, K. S., Jung, S. H., & Choy, H. E. (2005). DNA looping-mediated repression by histone-like protein H-NS: specific requirement of E σ 70 as a cofactor for looping. *Genes & Development*, 19(19), 2388.

Siryaporn, A., & Goulian, M. (2008). Cross-talk suppression between the CpxA–CpxR and EnvZ–OmpR two-component systems in *E. coli*. *Molecular Microbiology*, 70(2), 494–506.

Sorek, R., Zhu, Y., Creevey, C. J., Francino, M. P., Bork, P., & Rubin, E. M. (2007). Genome-wide experimental determination of barriers to horizontal gene transfer. *Science (New York, N.Y.)*, 318(5855), 1449–1452.

Sorokin, V., Severinov, K., & Gelfand, M. S. (2009). Systematic prediction of control proteins and their DNA binding sites. *Nucleic Acids Research*, 37(2), 441–451.

Stewart, A. J., & Plotkin, J. B. (2013). The evolution of complex gene regulation by low-specificity binding sites. *Proceedings. Biological Sciences*, 280(1768).

Tabor, S., & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of the National Academy of Sciences of the United States of America*, 82(4), 1074.

Tapias, A., López, G., & Ayora, S. (2000). Bacillus subtilis LrpC is a sequence-independent DNA-binding and DNA-bending protein which bridges DNA. *Nucleic Acids Research*, 28(2), 552.

Taylor, T. B., Mulley, G., Dills, A. H., Alsohim, A. S., McGuffin, L. J., Studholme, D. J., Silby, M. W., Brockhurst, M. A., Johnson, L. J., & Jackson, R. W. (2015). Evolution. Evolutionary resurrection of flagellar motility via rewiring of the nitrogen regulation system. *Science (New York, N.Y.)*, 347(6225), 1014–1017.

Taylor, T. B., Shepherd, M. J., Jackson, R. W., & Silby, M. W. (2022). Natural selection on crosstalk between gene regulatory networks facilitates bacterial adaptation to novel environments. In *Current Opinion in Microbiology* (Vol. 67, p. 102140). Elsevier Current Trends.

Thomas, C. M., & Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews. Microbiology*, 3(9), 711–721.

Tulchinsky, A. Y., Johnson, N. A., Watt, W. B., & Porter, A. H. (2014). Hybrid incompatibility arises in a sequence-based bioenergetic model of transcription factor binding. *Genetics*, 198(3), 1155–1166.

- Typas, A., & Hengge, R. (2006). Role of the spacer between the -35 and -10 regions in σ promoter selectivity in *Escherichia coli*. *Molecular Microbiology*, 59(3), 1037–1051.
- van Hijum, S. A. F. T., Medema, M. H., & Kuipers, O. P. (2009). Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. *Microbiology and Molecular Biology Reviews : MMBR*, 73(3), 481–509.
- Vasu, K., & Nagaraja, V. (2013). Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiology and Molecular Biology Reviews : MMBR*, 77(1), 53.
- Vijesurier, R. M., Carlock, L., Blumenthal, R. M., & Dunbar, J. C. (2000). Role and mechanism of action of C. PvuII, a regulatory protein conserved among restriction-modification systems. *Journal of Bacteriology*, 182(2), 477–487.
- Wagner, A. (2021). Adaptive gene misregulation. *Genetics*, 217(3).
- Wang, X., Kim, Y., Ma, Q., Hong, S. H., Pokusaeva, K., Sturino, J. M., & Wood, T. K. (2010). Cryptic prophages help bacteria cope with adverse environments. *Nature Communications* 2010 1:1, 1(1), 1–9.
- Wang, X., & Wood, T. K. (2016). Cryptic prophages as targets for drug development. *Drug Resistance Updates*, 27, 30–38.
- Wassarman, K. M., & Storz, G. (2000). 6S RNA regulates *E. coli* RNA polymerase activity. *Cell*, 101(6), 613–623.
- Willett, J. W., Tiwari, N., Müller, S., Hummels, K. R., Houtman, J. C. D., Fuentes, E. J., & Kirby, J. R. (2013). Specificity residues determine binding affinity for two-component signal transduction systems. *MBio*, 4(6).
- Wisniewska, A., Wons, E., Potrykus, K., Hinrichs, R., Gucwa, K., Graumann, P. L., & Mruk, I. (2022). Molecular basis for lethal cross-talk between two unrelated bacterial transcription factors - the regulatory protein of a restriction-modification system and the repressor of a defective prophage. *Nucleic Acids Research*, 50(19), 10964–10980.
- Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., Decruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z., & Graham, M. W. (1989). Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Research*, 17(9), 3469.
- Wunderlich, Z., & Mirny, L. A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. *Trends in Genetics : TIG*, 25(10), 434–440.
- Yan, K. K., Fang, G., Bhardwaj, N., Alexander, R. P., & Gerstein, M. (2010). Comparing genomes to computer operating systems in terms of the topology and evolution of their regulatory control networks. *Proceedings of the National Academy of Sciences of the United States of America*, 107(20), 9186–9191.

Yona, A. H., Alm, E. J., & Gore, J. (2018). Random sequences rapidly evolve into de novo promoters. *Nature Communications* 2018 9:1, 9(1), 1–10.

Zheng, D., Constantinidou, C., Hobman, J. L., & Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Research*, 32(19), 5874–5893.