



UNIwersytet
Warszawski

Wydział Biologii
Zakład Biologii Molekularnej
dr hab. Piotr Kozłowski



Warszawa, 9 sierpnia 2020 r.

Przedmiotem recenzji jest **praca doktorska "Charakterystyka kompleksu MLH1-MBD4, tworzonego przez białka biorące udział w naprawie DNA"** Pani mgr Katarzyny Poleszak, która jest efektem jej studiów doktoranckich w Instytucie Biochemii i Biofizyki PAN, ale została wykonana pod kierunkiem Prof. Janusza Bujnickiego w Laboratorium Bioinformatyki i Inżynierii Białka w Międzynarodowym Instytucie Biologii Molekularnej i Komórkowej, też w Warszawie.

MLH1 jest jednym z 4 ludzkich homologów bakteryjnego MutL i tak jak tamto, zaangażowane jest w naprawę DNA z udziałem koreplikacyjnego mechanizmu naprawy błędnie sparowanych zasad (MMR), zaś MBD4 jest glikozylazą usuwającą tyminę lub uracyl w parze z guaniną przy naprawie DNA z udziałem mechanizmu naprawy przez wycinanie (wadliwej) zasady (BER) i ma zdolność do wiązania metylowanych na cytozynie wysp CpG w DNA. Oba te białka oddziałują także z licznymi innymi białkami, w szczególności MLH1 z PMS2 - innym homologiem MutL, kiedy są razem zaangażowane w naprawę typu MMR oraz z Mre11 biorącym m.in. również udział w naprawie DNA, choć głównie przy uszkodzeniach dwuniciowych. Z licznymi mutacjami w genie *MLH1* (przynajmniej w części prowadzącymi do niesprawności systemu MMR poprzez zaburzenie interakcji MLH1-PMS2 i w efekcie powstawania nowych w innych genach) związany jest zespół Lyncha - dziedziczny rak jelita grubego, któremu mogą towarzyszyć nowotwory innych organów. Jednak, kompleks MLH1-MBD4 miałby być zaangażowany nie bezpośrednio w naprawę DNA, lecz w indukcję apoptozy w następstwie uszkodzeń w DNA. Zaburzenie oddziaływania między tymi dwoma białkami mogłoby więc i na tej drodze prowadzić do niekontrolowanych podziałów komórkowych, co jest podstawą procesu nowotworzenia.

Występowanie kompleksu MLH1-MBD4 stwierdzono już w 1999 r., kiedy także ustalono, że za tę interakcję odpowiada C-końcowy fragment MBR4 (aa 406-580) obejmujący domenę glikozylazy. Z kolei, w ciągu następnych 10 lat ustalono, że za interakcję MLH1 z innymi białkami odpowiada jego C-końcowa domena (CTD, aa 499-756), obejmująca subdomenę In i otaczającą ją subdomenę Ex, a w szczególności miejsce S2 w subdomenie In, które oddziałuje z silnie konserwowanym motywem MIP obecnym u części jego białkowych partnerów. Również, analizy Dr. Kosinskiego z grupy Prof. Bujnickiego wskazały na możliwy udział odpowiednika motywu QXLLXP w MLH1, także położonego w subdomenie In, w pobliżu miejsca S2, w oddziaływaniach tego białka z jego partnerami. Celem Doktorantki było zidentyfikowanie rejonów odpowiedzialnych za tworzenie kompleksu MLH1-MBD4, ustalenie jego budowy podjednostkowej, opracowanie jego modelu komputerowego i wyjaśnienie wpływu wybranych mutacji związanych z zespołem Lyncha na jego formowanie.

Doktorantka przedstawiła swoje badania w postaci rozprawy o klasycznym układzie, napisanej w języku polskim, zwięźle (ma 127 stron) i logicznie. Tekst uzupełniony jest 36 rysunkami, 28 tabelami, Wykazem skrótów i Literaturą obejmującą 155 pozycji. Wstęp dobrze wprowadza w metody użyte przez Doktorantkę do badania i modelowania kompleksu MLH1-MBD4, opisuje dwa mechanizmy naprawy jednoniciowych uszkodzeń DNA (MMR i BER) oraz nowotwory towarzyszące defektom w tych systemach naprawczych, a także podaje podstawową charakterystykę obu tych białek oraz ich kompleksu. Jednak, powielono w nim Rys. 1 (w miejsce Rys. 2) i zastosowano inny sposób cytowania w tekście (z et al.) i przy rysunkach i tabelach (z i in.). Ponadto, pomocnym byłby schemat dla naprawy BER (taki jak przedstawiono dla MMR) oraz ściślej zdefiniowanie (podanie sekwencji aminokwasowej i położenia) wszystkich przypuszczalnych rejonów oddziaływań, w szczególności MIP i S2. Materiały i metody podzielono na dwa rozdziały i są one przedstawione w stopniu szczegółowości umożliwiającym zrozumienie i ewentualne powtórzenie przeprowadzonych eksperymentów. Bardzo dogodnym jest zestawienie materiałów w postaci serii tabel. Te rozdziały mogą być doskonałą pomocą dla studentów dołączających do Zespołu i także dlatego należy przypomnieć, iż większość składowych genotypów dla szczepów bakteryjnych pisze się kursywą (Tab. 8), zaś nazwy plazmidów rozpoczyna się od litery "p" i członu wskazującego szkielet wektora (np. pET28-MBD4, str. 50). W pracy można też spotkać zwroty z

żargonu laboratoryjnego typu plazmidy "transformowano do" (str. 54) czy białko "wyekspresjonowano" (str. 77) lub nietypowe użycie terminu "dipeptyd" dla pary połączonych ze sobą peptydów (str. 81) oraz literówki, np. w sekwencji peptydu z MLH1 z odpowiednikiem motywu QXLLXP (powinno być 566-YDFANFGVLRLEPA-580, podkreślono literówki). Wspomniane nie tak liczne błędy nie stanowią jednak większej przeszkody dla merytorycznej oceny pracy.

W badaniach użyto szerokiego wachlarza zaawansowanych i uzupełniających się metod bioinformatycznych, biochemicznych i biofizycznych wraz z bardzo licznymi wariantami delecyjnymi oraz punktowymi (z substytucjami) MLH1 i MBD4 człowieka. Doktorantka wskazała poszukiwane rejony w obu badanych białkach (w MHL1: subdomenę Ex, odpowiednik motywu QXLLXP i miejsce S2, zaś w MBD4 m.in. sekwencję TSLYF, sklasyfikowaną przez nią jako MIP), ustaliła iż ich kompleks jest heterodimerem (dokładniej kompleks ich badanych *in vitro* fragmentów, odpowiednio aa 486-756 i aa 390-580) i opracowała w oparciu o uzyskane dane oraz rozwiązane przez innych struktury krystaliczne domeny glikozylazy ludzkiego MBD4 i C-końcowej domeny drożdżowego homologa MLH1, model komputerowy kompleksu tychże fragmentów obu białek. Nie udało się jednak wskazać oddziaływań pomiędzy konkretnymi resztami aminokwasowymi obu białek (punktów kontaktowych), chociaż dla kilku substytucji w genie *MLH1* związanych z zespołem Lyncha zaproponowano wyjaśnienie ich patogennego charakteru poprzez możliwy negatywny wpływ na tworzenie się kompleksu MLH1-MBD4. Wiarygodność techniczna tych wyników nie budzi zastrzeżeń i w planach jest ich opublikowanie. W opinii tego recenzenta, mimo ich obfitości, warto je jednak uzupełnić o badania kompleksu w żywych komórkach człowieka, np. zaburzeń w indukowanej kolokalizacji wariantów pełnej długości tych białek z substytucjami w postaci hybryd z białkami fluorescencyjnymi.

Mnogość tych wyników, ale także użycie odmiennych kombinacji substytucji MLH1 z ich badanej puli przy analizach różnymi metodami (możliwe było przebadanie wszystkimi 4 metodami tylko 2 z 15 tych wariantów, Wyniki 2), były zapewne powodem pominięcia omówienia substytucji F568A (Rys. 27) w tekście i Tab. 22 i 25. W takiej sytuacji pojawiają się także możliwe ich alternatywne interpretacje, zwłaszcza biorąc pod uwagę pewne ograniczenia użytych metod. Dla przykładu, mikromacierze peptydowe dobrze służą identyfikacji motywów sekwencyjnych, ale mogą nie wykrywać motywów przestrzennych, których punkty kontaktowe rozrzucone są po kilku peptydach i żaden z nich nie jest wystarczająco silny. Podobnie, użyty czynnik sieciujący BS³ należy do tych większych (szytywny łańcuch 8-węglowego nasyconego węglowodoru o długości 11,4 Å, a więc większy nawet od reszty najdłuższego aminokwasu białkowego - argininy) i dlatego może wiązać ze sobą raczej bardziej odległe od siebie reszty lizyn. Również, w opinii tego recenzenta, warto rozważyć alternatywną interpretację odnośnie motywu MIP i miejsca S2. Po pierwsze, wskazana sekwencja w MBD4 (TSLYF, aa 414-418) jako MIP odbiega znacząco od jego konsensusu (RSKY/CC), posiadającego reszty dwóch aminokwasów zasadowych (R i K), a których w niej całkowicie brak. Tak więc, nie negując znaczenia tych reszt aminokwasowych dla interakcji z MLH1 (potwierdzone wariantami z substytucjami każdej z tych reszt), warto rozważyć zmianę jego nazwy (np. na MIP typu 2). Po drugie, klasyczny MIP definiowany jest poprzez jego zdolność do wiązania się z miejscem S2 w MLH1 (aa 565-570 i 618-630). Jednak, uzyskane wyniki (Rys. 17 i 27) niekoniecznie to potwierdzają. MBD4 nie oddziałuje z dwoma peptydami z resztami aminokwasowymi tego miejsca (peptydy 606-620 i 616-630), w przeciwieństwie do Mre11 (innego partnera MHL1). Co prawda, MBD4 oddziałuje z trzecim peptydem z resztami z tego miejsca (peptyd 566-580), ale ten zawiera także odpowiednik motywu QXLLXP w MLH1 (LRLSEP, aa 574-579), który wcześniej przewidziano jako możliwe miejsce interakcji m.in. z MBD4, co Doktorantka sama potwierdziła wariantami z substytucjami w niektórych jego resztach. Przy analizie z pomocą wariantów MLH1 z substytucjami, tylko jedna (N570A) dotyczyła miejsca S2 i chociaż ten wariant wykazywał osłabioną interakcję z MBD4 to N570 jest tylko 3 reszty aminokwasowe przed tym odpowiednikiem QXLLXP, a poza tym, jest resztą tylko częściowo obecną na powierzchni białka (w przeciwieństwie do reszt tego odpowiednika) oraz substytucja N570A miałaby powodować destabilizację MLH1 (i prawie dwukrotny wzrost zawartości wstążek β), tak więc ta substytucja mogłaby wpływać na interakcję raczej pośrednio - wpływając na zachowanie pobliskiego odpowiednika motywu QXLLXP.

Reasumując, Doktorantka znacząco przybliżyła wyjaśnienie molekularnych podstaw formowania heterodimeru MLH1-MBD4, które mogą być pomocne m.in. przy diagnostyce, rokowaniu i projektowaniu nowych leków dla zespołu Lyncha. Recenzowana **praca spełnia wymagania** art. 16 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym ... (Dz.U. nr 65, poz. 595 z późn. zm.) i stąd **wnoszę o dopuszczenie** Pani mgr Katarzyny Poleszak do dalszych etapów przewodu doktorskiego.

Piotr Kozłowski



UNIwersytet
Warszawski

Wydział Biologii
Zakład Biologii Molekularnej
dr hab. Piotr Kozłowski



Warszawa, 9 sierpnia 2020 r.

Przedmiotem recenzji jest **praca doktorska "Charakterystyka kompleksu MLH1-MBD4, tworzonego przez białka biorące udział w naprawie DNA"** Pani mgr Katarzyny Poleszak, która jest efektem jej studiów doktoranckich w Instytucie Biochemii i Biofizyki PAN, ale została wykonana pod kierunkiem Prof. Janusza Bujnickiego w Laboratorium Bioinformatyki i Inżynierii Białka w Międzynarodowym Instytucie Biologii Molekularnej i Komórkowej, też w Warszawie.

MLH1 jest jednym z 4 ludzkich homologów bakteryjnego MutL i tak jak tamto, zaangażowane jest w naprawę DNA z udziałem koreplikacyjnego mechanizmu naprawy błędnie sparowanych zasad (MMR), zaś MBD4 jest glikozylazą usuwającą tyminę lub uracyl w parze z guaniną przy naprawie DNA z udziałem mechanizmu naprawy przez wycinanie (wadiowej) zasady (BER) i ma zdolność do wiązania metylowanych na cytozynie wysp CpG w DNA. Oba te białka oddziałują także z licznymi innymi białkami, w szczególności MLH1 z PMS2 - innym homologiem MutL, kiedy są razem zaangażowane w naprawę typu MMR oraz z Mre11 biorącym m.in. również udział w naprawie DNA, choć głównie przy uszkodzeniach dwuniciowych. Z licznymi mutacjami w genie *MLH1* (przynajmniej w części prowadzącymi do niesprawności systemu MMR poprzez zaburzenie interakcji MLH1-PMS2 i w efekcie powstawania nowych w innych genach) związany jest zespół Lyncha - dziedziczny rak jelita grubego, któremu mogą towarzyszyć nowotwory innych organów. Jednak, kompleks MLH1-MBD4 miałby być zaangażowany nie bezpośrednio w naprawę DNA, lecz w indukcję apoptozy w następstwie uszkodzeń w DNA. Zaburzenie oddziaływania między tymi dwoma białkami mogłoby więc i na tej drodze prowadzić do niekontrolowanych podziałów komórkowych, co jest podstawą procesu nowotworzenia.

Występowanie kompleksu MLH1-MBD4 stwierdzono już w 1999 r., kiedy także ustalono, że za tę interakcję odpowiada C-końcowy fragment MBR4 (aa 406-580) obejmujący domenę glikozylazy. Z kolei, w ciągu następnych 10 lat ustalono, że za interakcję MLH1 z innymi białkami odpowiada jego C-końcowa domena (CTD, aa 499-756), obejmująca subdomenę In i otaczającą ją subdomenę Ex, a w szczególności miejsce S2 w subdomenie In, które oddziałuje z silnie konserwowanym motywem MIP obecnym u części jego białkowych partnerów. Również, analizy Dr. Kosinskiego z grupy Prof. Bujnickiego wskazały na możliwy udział odpowiednika motywu QXLLXP w MLH1, także położonego w subdomenie In, w pobliżu miejsca S2, w oddziaływaniach tego białka z jego partnerami. Celem Doktorantki było zidentyfikowanie rejonów odpowiedzialnych za tworzenie kompleksu MLH1-MBD4, ustalenie jego budowy podjednostkowej, opracowanie jego modelu komputerowego i wyjaśnienie wpływu wybranych mutacji związanych z zespołem Lyncha na jego formowanie.

Doktorantka przedstawiła swoje badania w postaci rozprawy o klasycznym układzie, napisanej w języku polskim, zwięźle (ma 127 stron) i logicznie. Tekst uzupełniony jest 36 rysunkami, 28 tabelami, Wykazem skrótów i Literaturą obejmującą 155 pozycji. Wstęp dobrze wprowadza w metody użyte przez Doktorantkę do badania i modelowania kompleksu MLH1-MBD4, opisuje dwa mechanizmy naprawy jednoniciowych uszkodzeń DNA (MMR i BER) oraz nowotwory towarzyszące defektom w tych systemach naprawczych, a także podaje podstawową charakterystykę obu tych białek oraz ich kompleksu. Jednak, powielono w nim Rys. 1 (w miejsce Rys. 2) i zastosowano inny sposób cytowania w tekście (z et al.) i przy rysunkach i tabelach (z i in.). Ponadto, pomocnym byłby schemat dla naprawy BER (taki jak przedstawiono dla MMR) oraz ściślej zdefiniowanie (podanie sekwencji aminokwasowej i położenia) wszystkich przypuszczalnych rejonów oddziaływań, w szczególności MIP i S2. Materiały i metody podzielono na dwa rozdziały i są one przedstawione w stopniu szczegółowości umożliwiającym zrozumienie i ewentualne powtórzenie przeprowadzonych eksperymentów. Bardzo dogodnym jest zestawienie materiałów w postaci serii tabel. Te rozdziały mogą być doskonałą pomocą dla studentów dołączających do Zespołu i także dlatego należy przypomnieć, iż większość składowych genotypów dla szczepów bakteryjnych pisze się kursywą (Tab. 8), zaś nazwy plazmidów rozpoczyna się od litery "p" i członu wskazującego szkielet wektora (np. pET28-MBD4, str. 50). W pracy można też spotkać zwroty z

żargonu laboratoryjnego typu plazmidy "transformowano do" (str. 54) czy białko "wyekspresjonowano" (str. 77) lub nietypowe użycie terminu "dipeptyd" dla pary połączonych ze sobą peptydów (str. 81) oraz literówki, np. w sekwencji peptydu z MLH1 z odpowiednikiem motywu QXLLXP (powinno być 566-YDFANFGVLRLEPA-580, podkreślono literówki). Wspomniane nie tak liczne błędy nie stanowią jednak większej przeszkody dla merytorycznej oceny pracy.

W badaniach użyto szerokiego wachlarza zaawansowanych i uzupełniających się metod bioinformatycznych, biochemicznych i biofizycznych wraz z bardzo licznymi wariantami delecyjnymi oraz punktowymi (z substytucjami) MLH1 i MBD4 człowieka. Doktorantka wskazała poszukiwane rejony w obu badanych białkach (w MHL1: subdomenę Ex, odpowiednik motywu QXLLXP i miejsce S2, zaś w MBD4 m.in. sekwencję TSLYF, sklasyfikowaną przez nią jako MIP), ustaliła iż ich kompleks jest heterodimerem (dokładniej kompleks ich badanych *in vitro* fragmentów, odpowiednio aa 486-756 i aa 390-580) i opracowała w oparciu o uzyskane dane oraz rozwiązane przez innych struktury krystaliczne domeny glikozylazy ludzkiego MBD4 i C-końcowej domeny drożdżowego homologa MLH1, model komputerowy kompleksu tychże fragmentów obu białek. Nie udało się jednak wskazać oddziaływań pomiędzy konkretnymi resztami aminokwasowymi obu białek (punktów kontaktowych), chociaż dla kilku substytucji w genie *MLH1* związanych z zespołem Lyncha zaproponowano wyjaśnienie ich patogenicznego charakteru poprzez możliwy negatywny wpływ na tworzenie się kompleksu MLH1-MBD4. Wiarygodność techniczna tych wyników nie budzi zastrzeżeń i w planach jest ich opublikowanie. W opinii tego recenzenta, mimo ich obfitości, warto je jednak uzupełnić o badania kompleksu w żywych komórkach człowieka, np. zaburzeń w indukowanej kolokalizacji wariantów pełnej długości tych białek z substytucjami w postaci hybryd z białkami fluorescencyjnymi.

Mnogość tych wyników, ale także użycie odmiennych kombinacji substytucji MLH1 z ich badanej puli przy analizach różnymi metodami (możliwe było przebadanie wszystkimi 4 metodami tylko 2 z 15 tych wariantów, Wyniki 2), były zapewne powodem pominięcia omówienia substytucji F568A (Rys. 27) w tekście i Tab. 22 i 25. W takiej sytuacji pojawiają się także możliwe ich alternatywne interpretacje, zwłaszcza biorąc pod uwagę pewne ograniczenia użytych metod. Dla przykładu, mikromacierze peptydowe dobrze służą identyfikacji motywów sekwencyjnych, ale mogą nie wykrywać motywów przestrzennych, których punkty kontaktowe rozrzucone są po kilku peptydach i żaden z nich nie jest wystarczająco silny. Podobnie, użyty czynnik sieciujący BS³ należy do tych większych (sztywny łańcuch 8-węglowego nasyconego węglowodoru o długości 11,4 Å, a więc większy nawet od reszty najdłuższego aminokwasu białkowego - argininy) i dlatego może wiązać ze sobą raczej bardziej odległe od siebie reszty lizyn. Również, w opinii tego recenzenta, warto rozważyć alternatywną interpretację odnośnie motywu MIP i miejsca S2. Po pierwsze, wskazana sekwencja w MBD4 (TSLYF, aa 414-418) jako MIP odbiega znacząco od jego konsensusu (RSKY/CC), posiadającego reszty dwóch aminokwasów zasadowych (R i K), a których w niej całkowicie brak. Tak więc, nie negując znaczenia tych reszt aminokwasowych dla interakcji z MLH1 (potwierdzone wariantami z substytucjami każdej z tych reszt), warto rozważyć zmianę jego nazwy (np. na MIP typu 2). Po drugie, klasyczny MIP definiowany jest poprzez jego zdolność do wiązania się z miejscem S2 w MLH1 (aa 565-570 i 618-630). Jednak, uzyskane wyniki (Rys. 17 i 27) niekoniecznie to potwierdzają. MBD4 nie oddziałuje z dwoma peptydami z resztami aminokwasowymi tego miejsca (peptydy 606-620 i 616-630), w przeciwieństwie do Mre11 (innego partnera MHL1). Co prawda, MBD4 oddziałuje z trzecim peptydem z resztami z tego miejsca (peptyd 566-580), ale ten zawiera także odpowiednik motywu QXLLXP w MLH1 (LRLSEP, aa 574-579), który wcześniej przewidziano jako możliwe miejsce interakcji m.in. z MBD4, co Doktorantka sama potwierdziła wariantami z substytucjami w niektórych jego resztach. Przy analizie z pomocą wariantów MLH1 z substytucjami, tylko jedna (N570A) dotyczyła miejsca S2 i chociaż ten wariant wykazywał osłabioną interakcję z MBD4 to N570 jest tylko 3 reszty aminokwasowe przed tym odpowiednikiem QXLLXP, a poza tym, jest resztą tylko częściowo obecną na powierzchni białka (w przeciwieństwie do reszt tego odpowiednika) oraz substytucja N570A miałaby powodować destabilizację MLH1 (i prawie dwukrotny wzrost zawartości wstążek β), tak więc ta substytucja mogłaby wpływać na interakcję raczej pośrednio - wpływając na zachowanie pobliskiego odpowiednika motywu QXLLXP.

Reasumując, Doktorantka znacząco przybliżyła wyjaśnienie molekularnych podstaw formowania heterodimeru MLH1-MBD4, które mogą być pomocne m.in. przy diagnostyce, rokowaniu i projektowaniu nowych leków dla zespołu Lyncha. Recenzowana **praca spełnia wymagania** art. 16 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym ... (Dz.U. nr 65, poz. 595 z późn. zm.) i stąd **wnoszę o dopuszczenie** Pani mgr Katarzyny Poleszak do dalszych etapów przewodu doktorskiego.

Piotr Kozłowski