

395495 01

Peptyd o aktywności enzymatycznej białka typu Dicer, sposób otrzymywania krótkich cząsteczek RNA oraz jego zastosowanie

Przedmiotem wynalazku jest peptyd o aktywności enzymatycznej białka typu Dicer, sposób otrzymywania krótkich cząsteczek RNA oraz jego zastosowanie.

Organizmy eukariotyczne (rośliny i zwierzęta włącznie z człowiekiem) posiadają zdolność do generowania krótkich cząsteczek RNA o długości około 20-25 nukleotydów uczestniczących w regulacji ekspresji genów. Regulacja ekspresji genów przez krótkie cząsteczki RNA występuje w wielu ważnych procesach fizjologicznych (prolifracja i różnicowanie komórek, programowana śmierć komórki), jak i patologicznych (kancerogeneza, infekcje wirusowe, procesy neurodegeneracyjne). Do powstawania krótkich cząsteczek RNA niezbędny jest specyficzny enzym – białko wykazujące podobieństwo do RNazy typu III. Białko takie, w zależności od pochodzenia może nosić różne nazwy, w przypadku człowieka zwane jest Dicer, a w przypadku roślin – białkiem typu Dicer (DCL, ang. *Dicer-like protein*).

Większość białek typu Dicer (Dicer oraz DCL) występujących u kręgowców, owadów i roślin posiada w swej strukturze 6 typów domen: kasety DEAD, helikazy C, DUF283 (ang. *domain of unknown function*), PAZ (*Piwi/Argonaute/Zwill*), RNazy III oraz RBD (ang. *dsRNA binding domain*) [Margis et al.]. U niższych organizmów eukariotycznych białka z rodziny Dicer pozbawione są jednej bądź kilku z wymienionych domen. Przykładowo białko Dicer z pierwotniaka *Giardia intestinalis* zawiera jedynie domeny PAZ oraz RNazy III [Macrae et al.]. Wskazuje to na kluczową rolę tych dwóch domen dla aktywności katalitycznej białek typu Dicer.

Działanie białka typu Dicer polega na wycięciu krótkiego około 20-25 nukleotydowego dupleksu RNA z większej cząsteczki prekursorowej. Aby białko typu

Dicer właściwie spełniało swą rolę musi być zdolne do rozpoznania dwuniciowego regionu, z którego ma zostać wycięty krótki dsRNA oraz do bardzo precyzyjnego cięcia, tak by otrzymana cząsteczka spełniała ściśle określone parametry. Istotna jest nie tylko długość dupleksu RNA, ale i jego struktura. Musi on posiadać dwa niesparowane, wolne nukleotydy na końcu 3'. Należy podkreślić, iż cząsteczki dsRNA nie spełniające tych kryteriów nie będą skutecznie włączane do kompleksu efektorowego (RISC; ang. *RNA-induced silencing complex*), który uczestniczy w regulacji ekspresji genów. Po włączeniu krótkiego dupleksu do RISC jedna z nici RNA jest usuwana i degradowana, podczas gdy druga służy jako specyficzna sonda umożliwiająca rozpoznanie komplementarnego RNA lub DNA (genu).

W ostatnich latach krótkie cząsteczki RNA znajdują coraz szersze zastosowanie zarówno w biotechnologii jak i medycynie. Techniki wykorzystujące krótkie cząsteczki RNA do regulacji ekspresji genów są stosowane zarówno w celach poznawczych (np. do badania funkcji genów) oraz praktycznych (do uzyskania korzystnych pod względem użytkowym cech u organizmów roślinnych i zwierzęcych). Dodatkowo opracowywane są nowe metody terapeutyczne bazujące na preparatach zawierających krótkie cząsteczki RNA. Większość z tych technik wymaga zastosowania białka Dicer lub DCL do otrzymania krótkich dsRNA. Obecnie w skład komercyjnych zestawów stosowanych w badaniach aktywności biologicznej krótkich regulatorowych RNA wchodzi między innymi ekstrakt białkowy wzbogacony w Dicer ludzki lub z *Giardia intestinalis*.

Odnosnie do istniejących już patentów dotyczących zjawisk związanych z RNAi większość z nich dotyczy ludzkiego białka DICER – różniącego się od niniejszego wynalazku znacznie na poziomie sekwencji aminokwasowej, a także zastosowania sztucznych transgenów – zawierających krótkie sekwencje, kodujące cząsteczki konkretnych RNAi, skierowane do ściśle określonych genów – w celu transformacji roślin i modulacji ich fenotypu.

W zgłoszeniu patentowym WO2009117513 (opublikowanym 2009-09-24) opisano zmodyfikowany polipeptyd Dicer, który wykazuje wzmocnioną aktywność katalityczną. Rozwiązanie dostarcza także sposób otrzymywania małych regulatorowych RNA z dsRNA, obejmujący kontakt dsRNA z przedmiotowym modyfikowanym Dicer.

W zgłoszeniu patentowym US20100058490 (opublikowanym 2010-03-04) opisano sposoby wyciszania genów. W rozwiązaniu przedstawiono sposoby i środki do

modulowania wyciszenia genu u eukariota poprzez zmianę poziomu funkcjonalnego białka DICER i białek typu DICER. W rozwiązaniu przedstawiono także sposoby i środki do modulowania po-transkrypcyjnego wyciszania genu u eukariota poprzez zmianę poziomu funkcjonalnego białek zaangażowanych w wyciszanie transkrypcyjnego genu kodującego wyciszany RNA.

Pomimo istniejących do tej pory rozwiązań wykorzystujących krótkie cząsteczki RNA do regulacji ekspresji genów stosowanych zarówno do badania funkcji genów, jak i do uzyskania między innymi korzystnych pod względem użytkowym cech u organizmów roślinnych i zwierzęcych istnieje ciągła potrzeba wytwarzania krótkich cząsteczek RNA, o aktywności białka Dicer.

Celem przedmiotowego rozwiązania było opracowanie nowego sposobu wytwarzania krótkich cząsteczek RNA, wykorzystującego nowy, zaprojektowany przez twórców peptyd MtDCL1pepA o aktywności białka Dicer.

Realizacja tak określonego celu i rozwiązanie opisanych w stanie techniki problemów związanych z opracowaniem i dostarczeniem peptydu o aktywności Dicer, odróżniającego go od występujących w dostępnych preparatach pod względem pochodzenia oraz zoptymalizowanych parametrów fizykochemicznych i biochemicznych, rozpuszczalnego w roztworach wodnych, zostały osiągnięte w niniejszym wynalazku.

Wskazane powyżej cechy peptydu MtDCL1pepA przekładają się na szereg zalet proponowanej metody otrzymywania krótkich regulatorowych RNA. Proponowana metoda może być zdecydowanie tańsza od innych obecnie stosowanych, gdyż białko MtDCL1pepA można produkować zarówno w systemie eukariotycznym, jak i, co jest unikatową właściwością peptydu MtDCL1pepA, w tanim i niezwykle wydajnym systemie prokariotycznym. System ten pozwala także uzyskiwać preparat o niezwykle wysokiej czystości, zdecydowanie przewyższającej pozostałe opisane dotąd preparaty. Proponowana metoda, z racji wykorzystania enzymu roślinnego, umożliwia szczegółowe badania zjawiska interferencji RNA u roślin – jak dotąd nie istnieje możliwość produkcji krótkich regulatorowych RNA przy użyciu komercyjnego enzymu pochodzenia roślinnego.

Przedmiotem wynalazku jest peptyd, charakteryzujący się tym, że obejmuje peptyd MtDCL1pepA określony sekwencją SEQ. nr 1 o aktywności enzymatycznej białka typu Dicer.

Korzystnie, gdy zawiera peptyd MtDCL1pepA jego istotny fragment obejmujący co najmniej połowę sekwencji MtDCL1pepA lub sekwencje na poziomie podobieństwa sekwencji aminokwasowej co najmniej 80%.

Korzystnie, gdy zawiera peptyd MtDCL1pepA zawierający jedynie wybrane domeny niezbędne do zachowania własnej aktywności katalitycznej, i nie posiadającego 1154 aminokwasów z końca N białka MtDCL1, i że peptyd MtDCL1pepA zawiera domenę PAZ, dwie domeny RNazy III oraz dwie domeny RBD.

Korzystnie, gdy peptyd MtDCL1pepA zaopatrzony jest w znaczniki.

Korzystnie, gdy znaczniki obejmują peptyd S-transferazy glutationowej (GST) na N-końcu MtDCL1pepA i dwa znaczniki FLAG oraz heksahistydynowy (His) na C-końcu.

Korzystnie, gdy peptyd MtDCL1pepA jest wytwarzany w systemie eukariotycznym lub prokariotycznym.

Kolejnym przykładem wynalazku jest sposób otrzymywania krótkich cząsteczek RNA, charakteryzujący się tym, że stosuje się peptyd określony powyżej i, że sposób obejmuje:

- a) otrzymywanie cDNA kodującego peptyd MtDCL1 z rośliny *Medicago truncatula*;
- b) rekonstrukcje przypuszczalnej sekwencji cDNA (egzonów) genu kodującego MtDCL1;
- c) zaprojektowanie oligomerów DNA J08-10 określonego sekwencją SEQ. nr 2 i J08-13 określonego sekwencją SEQ. nr 3 oskrzydających sekwencje kodującą białko MtDCL1;
- d) przeprowadzanie reakcji syntezy pierwszej nici cDNA przy zastosowaniu RNA pochodzącego z młodych liści i młodych, szczytowych części naziemnych pędów *Medicago truncatula*;
- e) syntezę drugiej nici cDNA i amplifikację cDNA
- f) klonowanie cDNA kodującego MtDCL1, przy czym cDNA wprowadza się do wektora plazmidowego, przeprowadza się transformację bakterii i prowadzi hodowlę tych posiadających insert o długości 5500 pz
- g) wyizolowanie DNA plazmidowego do sekwencjonowania
- h) uzyskanie sekwencji białka MtDCL1 i poddanie jej analizie bioinformatycznej pod kątem zawartości znanych domen funkcjonalnych, charakterystycznych dla większości białek typu Dicer: kasety DEAD, helikaza C, DUF283, PAZ, RNaza III oraz RBD;

- i) zaprojektowanie skróconego białka MtDCL1pepA zawierającego jedynie wybrane domeny niezbędne do zachowania własnej aktywności katalitycznej, i nie posiadającego 1154 aminokwasów z końca N białka MtDCL1, przy czym usunięty fragment obejmuje kasetę DEAD, domenę helikazy C oraz DUF283, zaś peptyd MtDCL1pepA zawiera domenę PAZ, dwie domeny RNazy III oraz dwie domeny RBD.

Korzystnie, gdy peptyd MtDCL1pepA zaopatruje się w znaczniki.

Korzystnie, gdy znaczniki obejmują peptyd S-transferazy glutationowej (GST) na N-końcu MtDCL1pepA i dwa znaczniki FLAG oraz heksahistydynowy (His) na C-końcu.

Kolejnym przedmiotem wynalazku jest zastosowanie peptydu określonego powyżej do generowania krótkich 15-30 nukleotydowych cząsteczek RNA.

Rozwiązanie przedstawiono na rysunku, gdzie:

Figura 1 przedstawia amplifikację cDNA kodującego peptyd MtDCL1. Linie 1, 2, 3 – produkt PCR otrzymany w efekcie amplifikacji cDNA z młodych liści i szczytowej części pędu nadziemnego *M. truncatula* z zastosowaniem oligomerów DNA: J08-10 oraz J08-13. Wielkość produktu zgadza się z oczekiwaną dla cDNA DCL1 z *M. truncatula* – 5784 par zasad, oszacowaną na podstawie analiz bioinformatycznych cDNA dla DCL1 z innych organizmów oraz sekwencji klonu genomowego z *M. truncatula* - mth2-71o19, nr akcesyjny AC150443.

Figura 2 przedstawia strukturę klonu 44-57 kodującego peptyd DCL1 z *M. truncatula*. Porównanie sekwencji cDNA – klonu 44-57 z sekwencją klonu genomowego mth2-71o19 (nr akcesyjny AC150443) uwidacznia granice pomiędzy egzonami zaznaczone jako pionowe czarne linie; egzon 1, 14 i 17 zostały podpisane, a egzony 1, 2, 4, 6, 8, 10, 12, 14, 16, 18 i 20 zaznaczono jako prostokąty; granice pomiędzy egzonami oznaczono według koordynat odnoszących się do sekwencji klonu 44-57. Część kodująca klonu 44-57, oznaczona jako CDS (coding sequence) i zaznaczona na zielono, leży pomiędzy pozycją 79 a 5742 klonu 44-57. Pokazano też położenie regionów kodujących poszczególne domeny białkowe w peptydzie MtDCL1 44-57 kodowanym przez klon 44-57 i ich lokalizację w stosunku do granic pomiędzy egzonami.

Struktura sekwencji klonu 44-57 uwidacznia, iż klon ten zawiera kompletną sekwencję kodującą peptyd DCL1 – powyżej startu translacji, pozycji 79, znajduje się 78 nukleotydów a poniżej kodonu stop, pozycji 5742, jest 42 nukleotydowy odcinek nie

kodujący. Region kodujący (CDS) klonu 44-57 - leżący pomiędzy pozycjami 79 a 5742 - zawiera wszystkie domeny charakterystyczne dla peptydów DCL. W obrębie egzonu 13 i 15 widoczne jest niższe podobieństwo pomiędzy cDNA – sekwencją klonu 44-57 a sekwencją genomową pochodzącą z klonu mth2-71o19 (nr akcesyjny AC150443). Analizy dokonano przy pomocy programu Blastn.

Figura 3 przedstawia drzewo filogenetyczne peptydów DCL z *M. truncatula* i *A. thaliana*. Drzewo otrzymano metodą N-J (Neighbourhood-Joining) na podstawie sekwencji uporządkowanych programem ClustalW. W roślinie modelowej *A. thaliana* obecne są geny dla czterech typów białek DCL, nazywane: DCL1, DCL2, DCL3 i DCL4. Niektóre z tych białek występują w kilku formach splicingowych – oznaczonych na rysunku jako „-sf”. Cztery typy białek DCL obecne u *A. thaliana* mają podobny molekularny mechanizm działania ale różnią się funkcją. Przypuszcza się, że u innych roślin również występują geny dla wszystkich czterech typów białek DCL, tak jak u *A. thaliana*, przy czym niekiedy mogło dojść do duplikacji niektórych genów przez co doszło do powstania blisko spokrewnionych grup nazywanych rodzinami genowymi. Analiza filogenetyczna peptydu badanego pochodzącego z *M. truncatula* – peptydu kodowanego przez klon 44-57 – wobec wszystkich peptydów DCL z *A. thaliana* przyporządkowuje peptyd badany do peptydu ortologicznego. Przedstawione drzewo filogenetyczne pokazuje, iż pokrewieństwo pomiędzy peptydem MtDCL1 kodowany przez klon 44-57 a peptydem DCL1 z *A. thaliana* jest większe niż pomiędzy peptydem DCL1 z *A. thaliana* a jakimkolwiek innym peptydem DCL z *A. thaliana*. Stąd wnosi się, iż peptydy MtDCL1 z *M. truncatula* i DCL1 z *A. thaliana* są to peptydy ortologiczne. W analizie uwzględniono też peptyd otrzymany w efekcie bioinformatycznej analizy sekwencji klonu genomowego z *M. truncatula* mt2-71o19 (nr akcesyjny AC150443). Analizy filogenetycznej dokonano zestawem programów dostępnych na stronie internetowej <http://align.genome.jp>. W przypadku peptydów z *A. thaliana* podane są nazwy jak DCL1, DCL2 oraz numery akcesyjne peptydów, dla peptydu pochodzącego z klonu genomowego mt2-71o19 (wyróżnionego przez jednokrotne podkreślenie) podano nr akcesyjny sekwencji nukleotydowej, peptyd kodowany przez sekwencje cDNA klonu 44-57 otrzymany przez autorów oznaczono MtDCL1 peptyd (wyróżniony przez dwukrotne podkreślenie).

Figura 4 przedstawia budowę domenową białek typu Dicer z *Medicago truncatula* i *Arabidopsis thaliana*, człowieka i pierwotniaka *Giardia intestinalis*. Zaznaczono także

aktywny fragment białka DCL1 z *M. truncatula* – MtDCL1pepA. Identyfikacji domen dokonano przy pomocy narzędzia EMBL-EBI InterProScan [x3];

Figura 5 przedstawia: (A) schemat budowy wektora ekspresyjnego pGEX-6P-3 (GE Healthcare); (B) schemat wektora ekspresyjnego pGEX-MtDCL1pepA, otrzymanego z wektora pGEX-6P-3 i sekwencji kodującej MtDCL1pepA, wykorzystanego do produkcji peptydu MtDCL1pepA w komórkach bakteryjnych. Na schematach wskazano m.in. lokalizację sekwencji kodujących białko MtDCL1pepA, znaczniki GST, FLAG i His, lokalizację promotora tac (Ptac), genu selekcyjnego i miejsc restrykcyjnych wykorzystywanych w procedurze wklonowania sekwencji MtDCL1pepA;

(C) schemat reakcji PCR wykorzystanej do amplifikacji DNA kodującego białko MtDCL1pepA. Wskazana została budowa starterów.

Figura 6 przedstawia: (A) wynik ekspresji plazmidu kodującego białko MtDCL1pepA zaopatrzonego w znaczniki GST na końcu N oraz FLAG i His na końcu C (masa całego białka fuzyjnego wynosiła 112 kDa). Na żelu 10% PAA z SDS dokonano rozdziału frakcji białek izolowanych z próbek hodowli bakteryjnej szczepu transformowanego wektorem ekspresyjnym. Próbkę pobrano tuż przed indukcją ekspresji (czas 0h) oraz 4 godziny po dodaniu induktora – IPTG. Równocześnie prowadzono kontrolną hodowlę bez dodatku induktora ekspresji. Żel wybarwiono barwnikiem Comassine Blue; (B) otrzymany preparat białkowy wzbogacony w białko DCL1pepA pozbawione znacznika GST na końcu N (masa białka: 86 kDa). Preparat był rozdzielony na żelu 10% PAA z SDS. Czystość preparatu ilustruje żel wybarwiony Comassine Blue, białko zostało zidentyfikowane przy pomocy techniki Western Blot; (C) porównanie aktywności białka DCL1pepA oraz komercyjnie dostępnego białka Dicer z *Giardia intestinalis*. W reakcji z prekursorem ludzkiego miRNA 33a (66nt) w obecności dwóch różnych roztworów buforujących (B1, B2) białko DCL1pepA generuje produkty o długości ok. 20-25nt. Porównywane rekombinowane białko Dicer z *Giardia intestinalis* generuje, w obecności identycznego substratu, produkty o większym zakresie długości, z przewagą fragmentów 36-37 nukleotydowych. Reakcje były prowadzone w zoptymalizowanym roztworze buforującym dołączonym do białka Dicer z *Giardia intestinalis* (oznaczenie B1) lub buforze 20mM Tris-HCl pH 7,5 z 250mM NaCl i 2,5mM MgCl₂ (oznaczenie B2). M – marker mas cząsteczkowych, K – kontrola (mieszanina reakcyjna bez dodatku enzymu), HA – substrat miRNA 33a poddany hydrolizie alkalicznej, h – czas prowadzenia reakcji w godzinach.

W celu lepszego zrozumienia wynalazku poniżej przedstawiono przykładowe rozwiązanie według wynalazku.

Przykład

W bazach sekwencji (GenBank) nie ma zdeponowanej sekwencji cDNA dla białka DCL1 z *M. truncatula* (MtDCL1). Dostępna jest jedynie sekwencja genu (złożona z intronów i egzonów) oraz sztuczna sekwencja cDNA otrzymana w wyniku bioinformatycznej obróbki sekwencji genowej. Poznana sekwencja cDNA MtDCL1 różni się nieznacznie od sztucznej sekwencji cDNA MtDCL1 otrzymanej w wyniku obróbki bioinformatycznej sekwencji genomowej.

cDNA kodujące peptyd DCL1 z rośliny *Medicago truncatula* (dalej nazywany jako MtDCL1) otrzymano przy pomocy techniki RT PCR i klonowania wykorzystującego homologię. W pierwszym etapie przeszukano bazę sekwencji *Medicago truncatula* w GenBank przy pomocy aminokwasowej sekwencji białka DCL1 z *Arabidopsis thaliana*, nr akcesyjny NP_171612.1 i programu tblastn. Wybrano do dalszej pracy region [119169-109079] sekwencji klonu mth2-71o19 z *Medicago truncatula* o numerze akcesyjnym AC150443, dla którego podobieństwo z sekwencją białka DCL1 z *Arabidopsis thaliana* (nr akcesyjny NP_171612.1) cechuje się najniższą wartością oczekiwaną.

Region [119169-109079] sekwencji klonu mth2-71o19 z *Medicago truncatula* o numerze akcesyjnym AC150443 posłużył do rekonstrukcji przypuszczalnej sekwencji cDNA zawierającej kompletną sekwencję kodującą białko MtDCL1. Rekonstrukcja przypuszczalnej sekwencji cDNA (egzonów) genu kodującego MtDCL1 została przeprowadzona na drodze porównania regionu [119169-109079] sekwencji klonu mth2-71o19 (nr akcesyjny AC150443) z sekwencją kodującą DCL1 z *Arabidopsis thaliana* o numerze akcesyjnym NM_099986 przy pomocy programu Spidey (www.ncbi.nlm.nih.gov/spidey), oraz poprzez porównanie sekwencji aminokwasowej otrzymanej w wyniku translacji sekwencji klonu mth2-71o19 (nr akcesyjny AC150443) z sekwencją białka DCL1 z *Arabidopsis thaliana*. Bioinformatycznej translacji sekwencji klonu mth2-71o19 (nr akcesyjny AC150443) dokonano przy pomocy programów z zestawu Sequence Manipulation Suite (<http://www.bioinformatics.org/sms2>). Przyjęto, iż region [119169-109079] sekwencji klonu mth2-71o19, nr akcesyjny AC150443 zawiera kompletną sekwencję kodującą białko DCL1 z *M. truncatula* oraz część lub całość regionów cDNA nieulegających

translacji (UTR). Następnie zaprojektowano dwa oligomery DNA – J08-10 i J08-13, oskrzydlające sekwencję kodującą białko MtDCL1, których sekwencja była w 100% identyczna z wybranymi odcinkami regionu [119169-109079] sekwencji klonu mth2-71o19 (nr akcesyjny AC150443). Oligomer DNA o nazwie J08-10 składał się z 29 nukleotydów i miał sekwencję SEQ. nr 2: TAGAATAGGCGTTGATACACAGCAATAGG, podczas gdy oligomer J08-13 o sekwencji SEQ. nr 3: ACAACCACTGCTTGCTTCTGATTGG składał się z 25 nukleotydów (sekwencje podane zgodnie z konwencją od końca 5' do końca 3').

W kolejnym etapie prac przeprowadzono reakcję syntezy pierwszej nici cDNA stosując 2 mikrogramy RNA pochodzącego z młodych liści i młodych, szczytowych części nadziemnych pędów *Medicago truncatula* R108 na 20 mikrolitrów mieszaniny reakcyjnej oraz oligomer DNA (dT)18 o stężeniu końcowym 2,5 mikromoli / L, DTT o stężeniu końcowym 10 milimoli / L, dATP o stężeniu końcowym 0,5 milimoli / L, dCTP o stężeniu końcowym 0,5 milimoli / L, dGTP o stężeniu końcowym 0,5 milimoli / L, dTTP o stężeniu końcowym 0,5 milimoli / L, inhibitor RNAzy - RNaseOUT (Invitrogen) o stężeniu końcowym 2 jednostki / mikrolitr, oraz bufor do odwrotnej transkrypcji z zestawu SuperScript II Reverse Transcriptase (Invitrogen) i enzym – SuperScript II Reverse Transcriptase (Invitrogen) o stężeniu 10 jednostek / mikrolitr. Reakcję syntezy pierwszej nici cDNA prowadzono zgodnie z zaleceniami dostawcy zestawu SuperScript II Reverse Transcriptase (Invitrogen) z tym, iż inkubację w 42 st. C prowadzono przez 55 minut. Otrzymany w wyniku tej reakcji jednoniciowy cDNA użyto następnie, bez oczyszczania go od pozostałych składników reakcji odwrotnej transkrypcji, do syntezy drugiej nici cDNA i amplifikacji cDNA w reakcji PCR z zastosowaniem zestawu FastStart High Fidelity PCR System firmy Roche. Reakcję PCR prowadzono w buforze 2 (zawierającym chlorek magnezu o stężeniu końcowym w mieszaninie reakcyjnej 1,8 milimola / L) z zestawu FastStart High Fidelity PCR System (Roche), stosując 1 mikrolitr reakcji odwrotnej transkrypcji (opisanej wyżej) przy objętości końcowej mieszaniny reakcyjnej 50 mikrolitrów. W skład mieszaniny reakcyjnej wchodziło: DMSO o stężeniu końcowym 2%, dATP o stężeniu 0,2 milimola / L, dCTP o stężeniu 0,2 milimola / L, dGTP o stężeniu 0,2 milimola / L, dTTP o stężeniu 0,2 milimola / L, oligomer DNA J08-10 (sekwencja patrz wyżej) o stężeniu 0,3 mikromola / L, oligomer DNA J08-13 (sekwencja patrz wyżej) o stężeniu 0,3 mikromola / L, oraz mieszanina enzymów z zestawu FastStart High Fidelity PCR System (Roche) o stężeniu końcowym 0,05 jednostki / mikrolitr. Reakcję PCR

proszono stosując następujący program: etap pierwszy - inkubacja w 94 st. C przez 2 minuty, etap drugi: dziesięć razy sekwencja inkubacji: inkubacja w 94 st. C przez 30 sekund, inkubacja w 53 st. C przez 30 sekund, inkubacja w 68 st. C przez 6 minut, etap trzeci: dwadzieścia pięć razy sekwencja inkubacji: inkubacja w 94 st. C przez 30 sekund, inkubacja w 55 st. C przez 30 sekund, inkubacja w 68 st. C przez 6 minut z wydłużeniem czasu inkubacji o 10 sekund przy każdym kolejnym cyklu, etap czwarty: jeden raz inkubacja w 68 st. C przez 7 minut zakończona schłodzeniem reakcji do 4 st. C. Otrzymano w efekcie produkt o długości około 5784 pz (par zasad), Fig. 1. Produkt reakcji PCR oczyszczono na 0,7% żelu agarozowym, z którego wycięto DNA o długości fragmentów około 5.784 pz i dokonano ekstrakcji DNA z żelu agarozowego zestawem QIAquick Gel Extraction Kit firmy Qiagen postępując zgodnie z instrukcją producenta zestawu. Otrzymano preparat DNA o długości cząsteczek około 5.784 pz i stężeniu około 15 ng / mikrolitr, który użyto do kolejnego etapu – klonowania cDNA kodującego MtDCL1. Najpierw dokonano wprowadzenia badanego DNA do wektora plazmidowego pCR-XL-TOPO firmy (Invitrogen). Reakcję prowadzono zgodnie z instrukcją dostarczaną przez producenta zestawu TOPO XL PCR Cloning Kit – firmę Invitrogen, stosując 0,5 mikrolitra wyżej opisanego preparatu zawierającego klonowany cDNA kodujący MtDCL1 i 2,5 mikrolitra mieszaniny zawierającej aktywowany plazmid i enzym z zestawu TOPO XL PCR Cloning Kit (Invitrogen). Po zakończeniu reakcji, w której cDNA jest włączany do wektora plazmidowego dając zrekombinowane plazmidy, przeprowadzono transformację bakterii One Shot TOP10 Electrocomp E. coli (Invitrogen) zgodnie z instrukcją producenta. Do transformacji użyto 40 mikrolitrów bakterii, 1 mikrolitr zrekombinowanego plazmidu, kuwety do elektroporacji o odległości elektrod -1mm oraz elektroporator Gen-Pulser firmy BioLabs. W celu podania impulsu elektrycznego zastosowano 1250 V, 25 mikroFaradów i 200 Ohmów. Po inkubacji mieszaniny transformacyjnej z 250 mikrolitrami pożywki SOC w 37 st. C przez 75 min., wysiano po 20 mikrolitrów i 200 mikrolitrów bakterii na szalki zawierające pożywkę stałą LB z kanamycyną o stężeniu 50 mikrogramów / ml jako czynnikiem selekcyjnym i inkubowano je w 37 st. C przez 20 godzin. Otrzymano i przebadano 19 kolonii, z których trzy posiadały insert o spodziewanej długości – około 5.784 pz. Przeprowadzono hodowlę tych klonów w 20 ml pożywki płynnej i wyizolowano – metodą lizy alkalicznej [Sambrook et al.] – DNA plazmidowy, który użyto do sekwencjonowania. Analiza sekwencji wykazała, iż tylko jeden klon nazwany 41-57 zawiera kompletną sekwencję kodującą peptyd DCL1 pochodzący z *Medicago*

truncatula (MtDCL1), Fig-2. Potwierdzenie, iż peptyd kodowany przez klon 44-57 jest odpowiednikiem *M. truncatula* peptydu DCL1 z *A. thaliana* uzyskano w efekcie analizy filogenetycznej – Fig. 3, oraz tabela Tab. 1.

Tabela 1. Porównanie peptydów DCL1 z *Medicago truncatula* – tj. peptydu MtDCL1 kodowanego klon 44-57 oraz peptydu otrzymanego z bioinformatycznej analizy klonu genomowego mth2-71o19, nr akcesyjny AC150443 z peptydami DCL1, DCL2, DCL3 i DCL4 z *Arabidopsis thaliana*.

Tabela 1. Podobieństwo pomiędzy peptydami DCL z *M. truncatula* i *A. thaliana*. [%]

	MtDCL1 klon 44- 57	mth2- 71o19 AC150443	AthDCL1 NP171612.1	AthDCL2 NP001078101	AthDCL3 ABF19799.1	AthDCL4 AAZ80387.1
MtDCL1 Klon 44-57	100%	99.57%	86.54%	45.73%	43.13%	41.39%
mth2-71o19 AC150443	99.57%	100%	86.98%	45.88%	43.27%	41.53%
AthDCL1 NP171612.1	86.54%	86.98%	100%	45.88%	43.70%	40.96%
AthDCL2 NP001078101	45.73%	45.88%	45.88%	100%	40.81%	40.81%
AthDCL3 ABF19799.1	43.13%	43.27%	43.70%	40.81%	100%	38.49%
AthDCL4 AAZ80387.1	41.39%	41.53%	40.96%	40.81%	38.49%	100%

Stopień podobieństwa między parą peptydów wrażono jako procent identycznych aminokwasów na odpowiadających sobie pozycjach porównywanych peptydów. Korelację peptydów przyporządkowującą odpowiadające sobie pozycje w poszczególnych peptydach przeprowadzono programem ClustalW. Peptydy uporządkowane przez program ClustalW, przed analizą stopnia podobieństwa, zostały poddane oczyszczaniu z pozycji o niskiej wiarygodności korelacji oraz z regionów nie mających odpowiedników we wszystkich porównywanych sekwencjach przy pomocy programu Gblocks. Analizę prowadzono przy pomocy pakietu programów dostępnych na stronach internetowych <http://www.phylogeny.fr> oraz <http://www.bioinformatics.org/sms2/>.

Peptydy pochodzące z *M. truncatula* - peptyd MtDCL1 kodowany przez klon 44-57 oraz peptyd otrzymany w efekcie bioinformatycznej analizy sekwencji klonu genomowego mth2-71o19 (nr akcesyjny AC150443) są prawie dwa razy bardziej podobne do peptydu DCL1 z *A. thaliana* niż do pozostałych peptydów DCL z *A. thaliana*.

Podobieństwo pomiędzy peptydami pochodzącymi z *M. truncatula* - peptydem MtDCL1 kodowany przez klon 44-57 oraz peptydem otrzymany w efekcie bioinformatycznej analizy sekwencji klonu genomowego mth2-71o19 (nr akcesyjny AC150443) jest prawie dwukrotnie wyższe (1.89 – 2.10) od podobieństwa z pozostałymi peptydami DCL. Świadczy to – podobnie jak wynik analiz filogenetycznych, iż peptyd DCL1 z *A. thaliana* jest bliżej spokrewniony z peptydami MtDCL1 i z peptydem otrzymany w efekcie bioinformatycznej analizy sekwencji klonu genomowego mth2-71o19, niż z pozostałymi peptydami DCL z *A. thaliana*.

Uzyskaną sekwencję białka MtDCL1 (wynik translacji sekwencji DNA genu obecnego w klonie 41-57) poddano analizie bioinformatycznej pod kątem zawartości znanych domen funkcjonalnych, wykorzystując narzędzie EMBL-EBI InterProScan (<http://www.ebi.ac.uk/Tools/InterProScan>). W zadanej sekwencji zidentyfikowanych zostało 6 typów domen charakterystycznych dla większości białek typu Dicer: kasetę DEAD, helikaza C, DUF283, PAZ, RNaza III oraz RBD. Ich szczegółowe rozmieszczenie w białku MtDCL1 przedstawia Figura 4. Na podstawie wspomnianej analizy zaprojektowano białko skrócone, tzn. zawierające jedynie wybrane domeny, niezbędne do zachowania właściwej aktywności katalitycznej. Skrócenie białka było zabiegiem niezbędnym, by móc produkować białko w tanim i wydajnym bakteryjnym systemie ekspresyjnym. Za wzorzec przy wyborze domen posłużyło białko Dicer z pierwotniaka *Giardia intestinalis*, które posiadając jedynie 2 typy domen: PAZ i RNazy III, jest aktywne katalitycznie. Zaprojektowane skrócone białko, zwane dalej MtDCL1pepA, nie posiada 1154 aminokwasów z końca N białka MtDCL1. Usunięty fragment obejmuje kasetę DEAD, domenę helikazy C oraz DUF283. Peptyd MtDCL1pepA zawiera domenę PAZ, dwie domeny RNazy III oraz dwie domeny RBD, co obrazuje Figura 4.

Dla podniesienia wydajności ekspresji oraz zapewnienia możliwości zastosowania łatwych i efektywnych metod identyfikacji oraz oczyszczania białka zdecydowano o zaopatrzeniu MtDCL1pepA w kilka znaczników. I tak na końcu N MtDCL1pepA dołączono duży peptyd S-transferazy glutationowej (GST, ang. *glutathione S-transferase*), natomiast na końcu C dwa krótkie znaczniki: FLAG oraz heksahistydynowy (His). By otrzymać tak zaprojektowane białko przygotowano wektor ekspresyjny pGEX-MtDCLpepA, pochodną komercyjnie dostępnego plazmidu pGEX-6P-3 (GE Healthcare) zawierającego sekwencję znacznika GST (por. Figura 5).

Do plazmidu pGEX-6P-3 wklonowano fragment genu MtDCL1, kodujący wybrany fragment białka (aminokwasy 1155-1887) wraz ze znacznikami FLAG i His.

DNA do wklonowania uzyskano w dwóch reakcjach PCR, z wykorzystaniem trzech odmiennych starterów: starter FWD zawierał miejsce cięcia enzymu EcoRI oraz fragment sekwencji komplementarnej do sekwencji kodującej koniec N zaprojektowanego białka MtDCL1pepA, starter REV1 zawierał fragment sekwencji komplementarnej do sekwencji kodującej koniec C MtDCL1pepA oraz część sekwencji kodującej znaczniki FLAG i His, natomiast starter REV2 zawierał część sekwencji kodującej znaczniki FLAG i His oraz miejsce cięcia enzymu SalI. Sekwencje starterów REV1 i REV2 częściowo się pokrywały, by umożliwić przeprowadzenie reakcji PCR z wykorzystaniem REV2 na matrycy produktu reakcji PCR ze starterem REV1 (zob. Figura 5). W pierwszej reakcji PCR użyto DNA klonu 41-57 oraz startery FWD i REV1, natomiast w drugiej – startery FWD i REV2 oraz oczyszczony przy użyciu zestawu QIAquick PCR Purification Kit (QIAGEN) produkt poprzedniej reakcji PCR. Skład każdej reakcji PCR (podano stężenia końcowe, objętość reakcji: 50 μ l): DNA klonu 41-57 (100ng/50 μ l), primer FWD (1 μ M), primer REV1 lub REV2 (1 μ M), mix dNTP (200 μ M), bufor reakcyjny z zestawu Promega *Pfu* DNA Polymerase (1 \times), enzym z zestawu Promega *Pfu* DNA Polymerase (1,25U/50 μ l), woda wolna od nukleaz. Program reakcji: etap I (temp. 95 $^{\circ}$ C – 2 min.), etap II (sekwencja: temp. 95 $^{\circ}$ C – 1 min., 60 $^{\circ}$ C – 30 sek., 72 $^{\circ}$ C – 4 min.) powtórzony 30 razy, etap III (temp. 72 $^{\circ}$ C – 5 min.). Produkt reakcji oczyszczono przy użyciu zestawu QIAquick PCR Purification Kit (QIAGEN) zgodnie z opisem producenta. Tak przygotowane DNA insertu oraz DNA wektora pGEX-6P-3 poddano reakcjom trawienia enzymami restrykcyjnymi EcoRI (Fermentas) i SalI (Fermentas). Skład reakcji trawienia (podano stężenia końcowe, objętość reakcji: 40 μ l): DNA insertu lub wektora (1 μ g/40 μ l), Fermentas Bufor O (1 \times), enzym Fermentas EcoRI (5U/40 μ l), enzym Fermentas SalI (5U/40 μ l), woda wolna od nukleaz. Reakcję prowadzono 4 godziny w temp. 37 $^{\circ}$ C. Produkt każdej reakcji oczyszczono przy użyciu zestawu QIAquick PCR Purification Kit (QIAGEN) zgodnie z opisem producenta. Oczyszczone produkty trawień poddano reakcji ligacji z użyciem enzymu T4 DNA Ligase (Promega). Skład reakcji ligacji (podano stężenia końcowe, objętość reakcji: 20 μ l): DNA wektora (100ng/20 μ l), DNA insertu (200ng/20 μ l), bufor z zestawu Promega T4 DNA Ligase (1 \times), enzym z zestawu Promega T4 DNA Ligase (2U/20 μ l), woda wolna od nukleaz. Reakcję prowadzono 16 godz. w temp. 4 $^{\circ}$ C. Produktem reakcji transformowano komórki kompetentne *E. Coli* DH5 α , w celu

selekcji i namnożenia poprawnie skonstruowanych plazmidów pGEX-MtDCL1pepA. Do 50 μ l zawiesiny komórek kompetentnych dodano 5 μ l produktu ligacji, bakterie delikatnie mieszano, inkubowano w temp. 4°C przez 45 min. Następnie poddano szokowi termicznemu przez inkubację zawiesiny w temp. 42°C przez 45 sek. i szybkie schłodzenie w 4°C. Do zawiesiny dodano 1ml płynnej pożywki LB i wytrząsano przez godzinę w 37°C z prędkością 225rpm. Zawiesinę rozprowadzono na dwóch szalkach ze stałą pożywką LB z ampicyliną. Szalki inkubowano przez 16 godz. w temp. 37°C. Wybrano 24 pojedyncze kolonie bakterii wyhodowane na stałej pożywce LB z ampicyliną, przeniesiono do 2ml płynnej pożywki LB zawierającej ampicylinę i wytrząsano przez 16 godz. w 37°C z prędkością 300rpm. Każdą hodowlę wirowano z prędkością 14000 rpm przez 1 min., roztwór dekantowano, a z zawierającego bakterie osadu izolowano plazmidy metodą lizy alkalicznej. Każdy plazmid oczyszczono przy użyciu zestawu QIAquick PCR Purification Kit (QIAGEN) zgodnie z opisem producenta i sekwencjonowano. W ten sposób wyselekcjonowano preparat zawierający poprawnie skonstruowany, oczyszczony plazmid pGEX-MtDCL1pepA.

W celu przeprowadzenia procedury ekspresji białka gotowy plazmid wykorzystano do transformacji komórek kompetentnych szczepu *E. coli* BL21 (szczep ekspresyjny). Do 50 μ l preparatu komórek kompetentnych dodano 5 μ l oczyszczonego plazmidu pGEX-MtDCL1pepA (2ng/ μ l), delikatnie mieszano i inkubowano w temp. 4°C przez 30 min. Następnie bakterie poddano szokowi termicznemu przez inkubację zawiesiny w 42°C przez 30 sek. i szybkie schłodzenie w 4°C. Do zawiesiny dodano 250 μ l pożywki SOC i wytrząsano przez godzinę w temp. 37°C z prędkością 225 rpm, po czym zawiesinę rozprowadzono na dwóch szalkach ze stałą pożywką LB zawierającą ampicylinę. Szalki inkubowano przez 16 godz. w temp 37°C. Spośród otrzymanych na szalkach kolonii wybrano jedną, która posłużyła do inicjacji kultury ekspresyjnej. Kolonię przeniesiono do 10 ml płynnej pożywki LB zawierającej ampicylinę, hodowlę wytrząsano przez 16 godz. w temp. 37°C z prędkością 300 rpm i wykorzystano do zaszczerpienia 1000 ml świeżej pożywki LB z ampicyliną. Prowadzono dalszą inkubację hodowli w identycznych warunkach. W momencie osiągnięcia przez zawiesinę bakterii gęstości optycznej OD₆₀₀~0,7 temperaturę hodowli obniżono do 18°C i indukowano ekspresję przez dodanie roztworu izopropyl- β -D-1-tiogalaktopiranozydu (IPTG) do stężenia końcowego 0,05 mM. Ekspresję prowadzono przez kolejne 16 godzin. Zawiesinę bakterii wirowano następnie z prędkością 5000 rpm w 4°C przez 15 min, roztwór dekantowano, a osad bakteryjny wykorzystano do izolacji białka.

W celu izolacji rekombinowanego białka MtDCL1pepA przeprowadzono ekstrakcję totalnej frakcji białek rozpuszczalnych z bakterii. Osad bakteryjny zawieszono w buforze do ekstrakcji (140 mM NaCl, 2,7mM KCl, 10mM Na₂HPO₄, 1,8 mM KH₂PO₄, 5m M DTT, 1× CelLytic, 0,1 mg/ml lizozym, 25 U/ml benzonaza, pH 7,3) stosując proporcję 5 ml buforu na każdy 1g osadu, wytrząsano w temp. 23°C przez 15 minut i wirowano z prędkością 15000 rpm. Otrzymany supernatant zawierający frakcję bakteryjnych białek rozpuszczalnych analizowano na 10% denaturującym żelu poliakrylamidowym (SDS-PAGE). Wynik analizy przedstawia Figura 6. Z supernatantu izolowano MtDCL1pepA metodą chromatografii powinowactwa do glutationu. Supernatant nałożono na kolumnę zawierającą 1ml upakowanego złoża Glutathione Sepharose 4 Fast Flow przygotowanego zgodnie z opisem producenta. Złoże przemyto kolejno 10 ml buforu do wiązania (140mM NaCl, 2,7 mM KCl, 10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, pH 7,3) i 10 ml buforu do cięcia proteazą PreScission (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7,5), a następnie zmieszano z 1 ml buforu do cięcia proteazą PreScission z dodatkiem proteazy PreScission (40U/ml) i inkubowano w temp. 4°C przez 16 godzin. Białko eluowano z kolumny 1 ml buforu do cięcia proteazą PreScission, a następnie dokonano wymiany buforu na 50mM Tris-HCl pH 7,5 przy użyciu filtra Millipore Amicon Ultra-0.5 10K, zgodnie z opisem producenta. Otrzymany preparat (1 ml) analizowano na żelu (SDS-PAGE) oraz techniką Western-Blot (Figura 6).

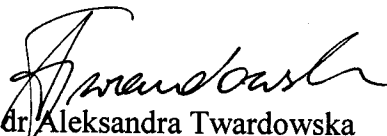
Aby określić aktywność otrzymanego peptydu przeprowadzono standardową reakcję trawienia prekursora miRNA (hsa-miR 33a) znakowanego radioaktywnie na końcu 5'. Dla porównania wykonano analogiczną serię reakcji trawienia, w których zamiast preparatu MtDCL1pepA użyto dostępne komercyjnie białko Dicer z *G. intestinalis*. Reakcje prowadzono w zoptymalizowanym komercyjnym buforze dołączonym do białka Dicer z *G. intestinalis*, a w przypadku peptydu MtDCL1pepA dodatkowo w buforze 20mM Tris-HCl pH 7,5 z 250 mM NaCl i 2,5 mM MgCl₂. We wszystkich przypadkach substrat reakcji (10 pikomoli) poddawano najpierw ogrzaniu w temp. 85°C przez 3 minuty, a następnie powolnemu schłodzeniu (1°C/min.) do temp. 23°C w celu uzyskania możliwie najbardziej homogennej struktury produktu. Po schłodzeniu do roztworu substratu dodawano właściwy bufor oraz enzym (preparat MtDCL1pepA – 7 µl, Dicer – zgodnie z opisem producenta). Reakcję prowadzono przez 16 godzin w temperaturze 37°C. Analizę produktów reakcji prowadzono metodą elektroforezy w 12% denaturującym żelu poliakrylamidowym (Figura 6). W reakcji z

peptydem MtDCL1pepA powstawał szereg produktów, z których większość mieści się w zakresie długości 20-25 nukleotydów, co odpowiada długości krótkich regulatorowych RNA. Reakcja z białkiem Dicer z *G. intestinalis* daje zdecydowanie inny zestaw produktów, z których dwa główne (36 oraz 37 nukleotydowe) są zdecydowanie dłuższe od typowych regulatorowych RNA (por. Figura 6).

Opisane powyżej wstępne testy aktywności dowiodły, że otrzymany peptyd MtDCL1pepA wykazuje oczekiwaną aktywność endorybonukleazy, katalizując reakcję wycięcia krótkich dupleksów RNA z dwuniciowego prekursora miRNA. Produkty te mają, zgodnie z oczekiwaniami, długości w zakresie 20-25 nukleotydów. Dowodzi to, że MtDCL1pepA posiada aktywność katalityczną charakterystyczną dla białek typu Dicer i może być z powodzeniem wykorzystany do produkcji małych regulatorowych RNA.

Zgłaszający: Instytut Chemii Bioorganicznej PAN

Pełnomocnik:


dr Aleksandra Twardowska
Rzecznik patentowy

Lista sekwencji

SEQ. nr 1

Sekwencja peptydu MtDCL1pepA

>MtDCL1pepA

DVVKASGLVPNRDSMETQNHINMTTKGKLMADTCTSPDDLVGRIVTAHSGKRFYVDS
 IRYEMTAENSFPRKEGYLGPLEYSSYADYYKQKYGVDLAYKQQPLIRGRGVPYCKNLLS
 PRFEHSEGHEDETEETHDKTYVFLPPELCLVPPLPGSLVRGAQRLPSIMRRVESMLLA
 VQLKNMINYPVQASKILEALTAASCQETFCYERAELLGDAYLKWVSRFLFLKHPQKHE
 GQLTRMRQQMVSNMVLYRYALSKGLQSYILADRFAPSRWAAPGVLPVFDEDTKDEESSL
 FDQERSIFKAERMDNTDEFEDEMEDGELESDSSSYRVLSSKTLADVVEALIGVYYVEGG
 KNAANHLMKWIGIHIEIDPEMECITRPSNVPDSILRSVDFDALEGALNIKFKDKGLLI
 ESITHASRPSSGVSCYQRLEFVGDAVLDHLITRHLFFSYTDLPPGRLTDLRAAAVNNEN
 FARVTVKHNHLHLRHHGSSALEKQIKDFVREVQDELSPGFNSFGLGDCKAPKVLGDIL
 ESIAGAIFLDSGRNTAVVWKFQPLLHPMVTPETLPMHPVRELQERCQQQAEGLEYRAS
 RAGNLATVEVFIDGVQVGAAQNPQKKMAQKLAARNALALKEKEESKIQEKNDEKETKS
 GNQTFTRQTLNDICLRRNWPMFPCYRCVSEGPAHAKKFTFAVRVNTTDKGWTDECVGEP
 MPSVKKAKDSA AVL L L L L L E L I N K L Y S S

SEQ. nr 2

oligomer J08-10

TAGAATAGGCGTTGATACACAGCAATAGG

SEQ. nr 3

oligomer J08-13

ACAACCACTGCTTGCTTCTGATTGG

Literatura:

Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM (2006) The evolution and diversification of Dicers in plants. *FEBS Lett* 580:2442–2450
Science. 2006 Jan 13;311(5758):195-8. Structural basis for double-stranded RNA processing by Dicer. Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA.

[x3] *The Pfam protein families database*: R.D. Finn, J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, O.L. Gavin, P. Guneseakaran, G. Ceric, K. Forslund, L. Holm, E.L. Sonnhammer, S.R. Eddy, A. Bateman **Nucleic Acids Research** (2010) Database Issue 38:D211-22

Sambrook J., Fritsch E., Maniatis T., *Molecular Cloning A Laboratory manual*, 1989, Second Edit., Cold Spring Harbor Lab. Press, pp.1.26-1.28.