

Instytut Chemii Bioorganicznej
Polskiej Akademii Nauk
Zakład Biologii Molekularnej i Systemowej

Lucyna Budźko

Otrzymywanie ludzkiej deaminazy cytydyny
indukowanej aktywacją limfocytów B oraz badania
mechanizmów molekularnych leżących u podłoża jej
aktywności biologicznej

Production of human activation-induced cytidine deaminase
and studies of molecular mechanisms underlying its biological activity

Promotor

prof. dr hab. Marek Figlerowicz

Promotor pomocniczy

dr Paulina Jackowiak

Poznań 2015

*Pracę tę dedykuję Mojej Rodzinie
w podziękowaniu za bezwarunkową miłość i wsparcie*

*Składam serdeczne podziękowania
Prof. dr hab. Markowi Figlerowiczowi
za kierownictwo naukowe, ogromną życzliwość
i poświęcony czas*

*Szczególne podziękowanie składam
dr Paulinie Jackowiak za przekazaną wiedzę,
cierpliwość i nieocenioną pomoc
oraz całemu zespołowi Zakładu Biologii Molekularnej
i Systemowej za wspaniałą atmosferę współpracy*

SPIS TREŚCI

Wykaz skrótów stosowanych w pracy.....	5
1. WPROWADZENIE	9
1.1. Rodzina białek AID/APOBEC.....	10
1.1.1. Ewolucja i organizacja genów.....	10
1.1.2. Funkcje	13
1.1.3. Struktura i właściwości biochemiczne	16
1.1.4. Podsumowanie	22
1.2. Odkrycie AID oraz roli jaką białko to odgrywa w procesie powstawania przeciwciał	23
1.3. Regulacja ekspresji genu AID.....	28
1.3.1. Regulacja transkrypcji.....	28
1.3.2. Regulacja potranskrypcyjna	31
1.3.3. Regulacja potranslacyjna.....	31
1.4. Specyficzność AID względem genów immunoglobulin.....	34
1.5. AID a aktywna demetylacja genomu	39
1.5.1. Metylacja i demetylacja genomu.....	39
1.5.2. Proponowane mechanizmy aktywnej demetylacji	40
1.5.3. Udział AID w procesach aktywnej demetylacji	43
1.6. Dysfunkcja AID	49
1.6.1. Udział AID w destabilizacji genomu	49
1.6.2. Udział AID w powstawaniu nowotworów	51
1.6.3. Zaburzenia ekspresji <i>AID</i> powodowane przez patogeny	53
2. CEL PRACY	57
3. MATERIAŁY I METODY	59
3.1. Wykaz materiałów stosowanych w pracy.....	59
3.1.1. Podstawowe odczynniki, złoże, enzymy oraz zestawy gotowych odczynników	59
3.1.2. Oligodeoksyrybonukleotydy	61
3.1.3. Plazmidy.....	62
3.1.4. Szczepy bakteryjne.....	63
3.1.5. Przeciwciała	64
3.1.6. Roztwory i bufory	64
3.1.7. Pożywki.....	69

3.1.8. Żele.....	71
3.1.9. Bufory obciążające i wzorce mas.....	72
3.2. Wykaz metod stosowanych w pracy.....	74
3.2.1. Metody stosowane w pracy z kwasami nukleinowymi.....	74
3.2.1.1. Łańcuchowa reakcja polimerazy (PCR).....	74
3.2.1.2. Oczyszczanie DNA.....	75
3.2.1.3. Rozdział elektroforetyczny kwasów nukleinowych w żelu agarozowym w warunkach natywnych.....	75
3.2.1.4. Rozdział elektroforetyczny kwasów nukleinowych w żelu poliakrylamidowym w warunkach denaturujących.....	76
3.2.1.5. Wizualizacja i oznaczanie stężenia kwasów nukleinowych.....	76
3.2.1.6. Sekwencjonowanie DNA.....	77
3.2.1.7. Trawienie enzymami restrykcyjnymi produktów PCR oraz plazmidów.....	78
3.2.1.8. Ligacja.....	79
3.2.1.9. Ligacja z wektorem pGEM-T Easy.....	79
3.2.1.10. Izolacja plazmidów w małej skali (do 10 µg).....	79
3.2.1.11. Izolacja plazmidów w średniej skali (do 100 µg).....	80
3.2.2. Metody rozdziału i analizy białek.....	80
3.2.2.1. Rozdział elektroforetyczny białek w żelu poliakrylamidowym w warunkach denaturujących (żel PAA).....	80
3.2.2.2. Identyfikacja białek metodą <i>western blot</i>	81
3.2.3. Otrzymywanie AID i jej mutantów w systemie prokariotycznym.....	82
3.2.3.1. Klonowanie genu AID do wektora ekspresyjnego.....	82
3.2.3.2. Otrzymywanie zmutowanych wariantów AID.....	82
3.2.3.3. Transformacja komórek bakteryjnych metodą szoku cieplnego.....	84
3.2.3.4. Wytwarzanie białka AID w komórkach bakteryjnych.....	84
3.2.3.5. Produkcja AID i jej mutantów w komórkach bakteryjnych w skali analitycznej i przygotowanie stoków glicerolowych.....	85
3.2.3.6. Analiza rozpuszczalności białek.....	86
3.2.3.7. Produkcja AID i jej mutantów w komórkach bakteryjnych w skali preparatywnej.....	87
3.2.3.8. Oczyszczanie białek otrzymanych w komórkach bakteryjnych.....	88
3.2.3.9. Wymiana buforu i zagęszczanie preparatów białkowych.....	90
3.2.3.10. Sączenie molekularne.....	90

3.2.3.11. Oznaczanie stężenia białka metodą Bradford.....	90
3.2.4. Metody analizy aktywności enzymatycznej AID.....	91
3.2.4.1. Radioizotopowe znakowanie DNA na końcu 5'	91
3.2.4.2. Badanie aktywności deaminazowej AID wobec pojedynczego motywu WRCY	92
3.2.4.3. Badanie aktywności deaminazowej AID wobec pojedynczego motywu WR5mCY.....	93
3.2.4.4. Badanie aktywności deaminazowej AID wobec pojedynczego motywu WR5hmCY.....	94
3.2.4.5. Badanie aktywności deaminazowej AID na matrycach zawierających wszystkie sekwencje typu NNCN.....	95
3.2.4.6. Badanie wpływu AID na globalny poziom 5mC w genomowym DNA	97
3.2.5. Modelowanie struktury trzeciorzędowej AID, dynamika molekularna w polu siłowym <i>Charmm36</i> , dokowanie substratów do centrum aktywnego AID ...	98
4. WYNIKI.....	100
4.1. Otrzymywanie białka hAID w systemie prokariotycznym.....	100
4.2. Wstępna charakterystyka otrzymanego białka.....	117
4.3. Wybór i otrzymywanie mutantów hAID	122
4.4. Charakterystyka aktywności hAID i jej mutantów w systemie <i>in vitro</i>	125
4.4.1. Charakterystyka aktywności deaminazowej hAID i jej mutantów wobec pojedynczych motywów: WRCY, WR5mCY, WR5hmCY	125
4.4.2. Charakterystyka aktywności deaminazowej hAID i jej mutantu N51A wobec matryc zawierających wszystkie sekwencje typu NNCN.....	129
4.4.3. Badanie wpływu hAID i jej mutantów na poziom 5mC w genomowym DNA <i>in vitro</i>	132
4.5. Bioinformatyczna analiza struktury hAID oraz jej kompleksów z substratami	134
5. DYSKUSJA.....	147
6. WNIOSKI.....	158
7. STRESZCZENIE	160
8. ABSTRACT	162
9. FINANSOWANIE	163
10. BIBLIOGRAFIA	164
11. ZAŁĄCZNIKI.....	174

Wykaz skrótów stosowanych w pracy

14-3-3 (ang. *tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide*) polipeptyd teta białka aktywującego 3-monooksygenazę tyrozyny/5-monooksygenazę tryptofanu

3'E (ang. *enhancer 3'*) wzmacniacz 3'

3'UTR (ang. *3' untranslated region*) 3' - końcowy region nieulegający translacji

AAV (ang. *Adeno-associated Virus*) wirus towarzyszący adenowirusom

ACF (ang. *APOBEC1 complementation factor*) czynnik komplementujący dla APOBEC1

ADAT2 (ang. *adenosine deaminase, tRNA-specific 2*) deaminaza adenozyiny 2, specyficzna wobec tRNA

AID (ang. *activation-induced cytidine deaminase*) deaminaza cytydyny indukowana aktywacją limfocytów B

Alkb (ang. *alpha-ketoglutarate-dependent dioxygenase*) dioksygenaza zależna od kwasu alfa-ketoglutazarowego

AP (ang. *aprimidinic site*) miejsce apirimidynowe

APE1 (ang. *aprimidinic endonuclease 1*) endonukleaza apirimidynowa 1

APE2 (ang. *aprimidinic endonuclease 2*) endonukleaza apirimidynowa 2

apoB - apolipoproteina B

APOBEC (ang. *apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like*) enzym edytujący mRNA apolipoproteiny B typu katalicznego polipeptyd

ATM (ang. *ataxia telangiectasia mutated*) gen zmutowany w zespole ataksja – teleangiektazja (koduje kinazę serynowo-treoninową)

B29, CD79b (ang. *cluster of differentiation 79, immunoglobulin-associated beta*) antygen różnicowania komórkowego 79 beta

BATF (ang. *basic leucine zipper transcription factor, ATF-like*) podstawowy czynnik transkrypcyjny typu ATF, zawierający motyw suwaka leucynowego

BCL-6 (ang. *B-cell lymphoma 6 protein*) białko 6 związane z chłoniakiem limfocytarnym B-komórkowym

B-CLL (ang. *B-cell chronic lymphocytic leukemia*) przewlekła białaczka limfatyczna B-komórkowa

BCR (ang. *B cell receptor*) receptor limfocytów B

BER (ang. *base excision repair*) mechanizm naprawy przez wycięcie zasad

BL (ang. *Burkitt Lymphoma*) chłoniak Burkitta

C/EBP (ang. *CCAAT/enhancer binding protein*) białko wiążące wzmacniacz CCAAT

CD (ang. *catalytic domain*) domena kataliczna

CD154, CD40LG (ang. *CD40 ligand*) ligand receptora CD40

CD40 (ang. *cluster of differentiation 40, TNF receptor superfamily member 5*) antygen różnicowania komórkowego 40

CD81 (ang. *cluster of differentiation 81*) antygen różnicowania komórkowego 81

CDA (ang. *cytidine deaminase*) deaminaza cytydyny

ChIP-seq (ang. *chromatin immunoprecipitation followed by massive parallel sequencing*) immunoprecypitacja chromatyny połączona z sekwencjonowaniem nowej generacji

CSR (ang. *class switch recombination*) rekombinacyjne przełączanie klas

- CTNBL1 (ang. *cadherin-associated protein, beta like 1*) podobne do beta 1 białko związane z kadherynami
- DSBs (ang. *double-stranded DNA breaks*) dwuniciowe pęknięcia DNA
- DEMETER (ang. *transcriptional activator DEMETER*) aktywator transkrypcyjny DEMETER
- DNA (ang. *deoxyribonucleic acid*) kwas deoksyrybonukleinowy
- DnaJ1 (ang. *DnaJ (Hsp40) homolog, subfamily A, member 1*) białko 1 podrodziny A homologów DnaJ (Hsp40)
- DNA-PK (ang. *DNA-dependent protein kinase*) kinaza białkowa zależna od DNA
- DNMT (ang. *DNA methyltransferase*) metylotransferaza DNA
- E2A (ang. *E2A transcription factor*) czynnik transkrypcyjny E2A
- E2F (ang. *E2F transcription factor*) czynnik transkrypcyjny E2F
- EBV (ang. *Epstein-Barr Virus*) wirus Epsteina-Barr
- eEF1A (ang. *Eukaryotic translation elongation factor 1 alpha*) eukariotyczny czynnik elongacji translacji 1A
- EIAV (ang. *Equine Infectious Anemia Virus*) wirus niedokrwistości zakaźnej koni
- ERE (ang. *estrogen response element*) element odpowiedzi na estrogen
- ES (ang. *embryonic stem cells*) embrionalne komórki macierzyste
- FAS, APO-1 (ang. *fas cell surface death receptor; apoptosis antigen 1*) antygen 1 apoptozy
- FL (ang. *Follicular Lymphoma*) chłoniak grudkowy
- FV (ang. *Foamy Virus*) wirus pianisty
- Gadd45 (ang. *growth arrest and DNA damage-inducible protein 45*) białko 45 indukowane zahamowaniem wzrostu i uszkodzeniami DNA
- GANP (ang. *germinal center-associated nuclear protein*) białko jądrowe związane z centrum namnażania limfocytów w guzłach limfatycznych
- GC (ang. *germinal center*) centrum namnażania limfocytów (in. ośrodek odczynowy) w guzłach limfatycznych
- GFP (ang. *green fluorescent protein*) białko zielonej fluorescencji
- HBV (ang. *Hepatitis B Virus*) wirus zapalenia wątroby typu B
- HCC (ang. *hepatocellular carcinoma*) pierwotny rak wątrobowokomórkowy
- HCV (ang. *Hepatitis C Virus*) wirus zapalenia wątroby typu C
- HIGM2 (ang. *hyper IgM syndrome 2*) hiperimmunoglobulinemia M typu 2
- HIGM5 (ang. *hyper IgM syndrome 5*) hiperimmunoglobulinemia M typu 5
- HIV (ang. *Human Immunodeficiency Virus*) ludzki wirus niedoboru odporności
- HoxC4 (ang. *homeobox C4*) kasetta homeo C4
- HPV (ang. *Human Papilloma Virus*) ludzki wirus brodawczaka
- Hsp40 (ang. *heat shock protein 40*) białko szoku cieplnego 40
- Hsp70 (ang. *heat shock protein 70*) białko szoku cieplnego 70
- Hsp90 (ang. *heat shock protein 90*) białko szoku cieplnego 90
- HSV (ang. *Herpes Simplex Virus*) wirus opryszczki pospolitej
- iE (ang. *intronic enhancer*) wzmacniacz intronowy
- IL13 (ang. *interleukin 13*) interleukina 13

- IL1 β (ang. *interleukin 1 β*) interleukina 1 β
- IL4 (ang. *interleukin 4*) interleukina 4
- iPSC (ang. *induced pluripotent stem cells*) indukowane pluripotencjalne komórki macierzyste
- KAP1 (ang. *KRAB-associated protein-1*) białko 1 towarzyszące białku KRAB
- LPS (ang. *lipopolysaccharide*) lipopolisacharyd
- MB, PSMB5 (ang. *proteasome subunit, beta type, 5*) podjednostka 5 proteasomu, typu beta
- MBD4 (ang. *methyl-CpG binding domain protein 4*) białko 4 z domeną wiążącą zmetylowane motywy CpG
- MCL (ang. *mantle cell lymphoma*) chłoniak z komórek płaszczka
- MHCC (ang. *metastatic hepatocellular carcinoma*) przerzutowy pierwotny rak wątrobowokomórkowy
- miRNA (ang. *microRNA*) mikro RNA
- MLV (ang. *Murine Leukemia Virus*) myszy wirus białaczki
- MM (ang. *multiple myeloma*) szpiczak mnogi
- MMR (ang. *mismatch repair*) mechanizm naprawy niesparowanych zasad
- mRNA (ang. *messenger RNA*) matrycowy (informacyjny) RNA
- MSH2 (ang. *mutS homolog 2*) homolog 2 białka mutS
- MSH6 (ang. *mutS homolog 6*) homolog 6 białka mutS
- MYB (ang. *v-myb avian myeloblastosis viral oncogene homolog*) homolog onkogenu v-myb wirusa ptasiej mieloblastozy
- MYC (ang. *v-myc avian myelocytomatosis viral oncogene homolog*) homolog onkogenu v-myc wirusa ptasiej białaczki
- NANOG (ang. *nanog homeobox protein*) białko homeokasety nanog
- Nbs1 (ang. *Nijmegen breakage syndrome protein 1*) białko 1 zespołu Nijmegen
- NER (ang. *nucleotide excision repair*) mechanizm naprawy mutacji przez wycięcie nukleotydów
- NES (ang. *nuclear export signal*) sygnał eksportu jądrowego
- NEUROD2 (ang. *neuronal differentiation 2*) czynnik 2 różnicowania neuronów
- NF κ B (ang. *nuclear factor kappa-light-chain-enhancer of activated B cells*) czynnik jądrowy kappa B
- NHEJ (ang. *non-homologous end joining*) niehomologiczne łączenie końców
- NLS (ang. *nuclear localization signal*) sygnał lokalizacji jądrowej
- NMR (ang. *nuclear magnetic resonance*) spektroskopia magnetycznego rezonansu jądrowego
- OCT4 (ang. *octamer-binding transcription factor 4*) czynnik transkrypcyjny 4 wiążący oktamer
- P19, CDKN2D (ang. *cyclin-dependent kinase inhibitor 2D*) inhibitor 2D kinazy zależnej od cyklin
- P53, TP53 (ang. *tumor protein p53*) supresor nowotworów p53
- PAX5 (ang. *paired box 5*) sparowana kasetta 5
- PGCs (ang. *primordial germ cells*) pierwotne komórki płciowe
- PIM1 (ang. *proviral insertion site in moloney murine leukemia virus 1*) onkogen/kinaza pim-1
- PKA (ang. *protein kinase A*) kinaza białkowa A
- PRE (ang. *progesterone binding element*) element wiążący progesteron
- PTBP2 (ang. *polypyrimidine tract binding protein 2*) białko 2 wiążące trakty polipirimidynowe
- RAS (ang. *rat sarcoma viral oncogene homolog*) homolog onkogenu szczurzego wirusa mięsaka

- RB1 (ang. *retinoblastoma 1*) białko 1 związane z siatkówczakiem
- RHOH (ang. *ras homolog family member H*) białko H rodziny homologów białek ras
- RNA (ang. *ribonucleic acid*) kwas rybonukleinowy
- ROS (ang. *repressor of silencing 1*) represor wyciszania 1
- RPA (ang. *replication protein A*) białko replikacyjne A
- SHM (ang. *somatic hypermutation*) hiperzmiennność somatyczna
- SIV (ang. *Simian Immunodeficiency Virus*) małpi wirus niedoboru odporności
- Smad3 (ang. *SMAD family member 3*) białko 3 rodziny SMAD
- Smad4 (ang. *SMAD family member 4*) białko 4 rodziny SMAD
- SMUG (ang. *single – strand selective monofunctional uracil-DNA glycosylase*) jednofunkcyjna glikozylaza uracylo-DNA selektywna wobec ssDNA
- Sp1 (ang. *specificity protein 1*) białko specyficzności 1
- Sp3 (ang. *specificity protein 3*) białko specyficzności 3
- Spt5 (ang. *suppressor of Ty (yeast transposable elements) 5; transcription elongation factor Spt5*) supresor Ty5, transkrypcyjny czynnik elongacyjny Spt5
- SRSF1 (ang. *serine/arginine-rich splicing factor 1*) czynnik 1 dojrzewania mRNA bogaty w seryny/argininy
- SSBs (ang. *single-strand DNA breaks*) jednoniciowe pęknięcia DNA
- ssDNA (ang. *single-stranded DNA*) jednoniciowy kwas deoksyrybonukleinowy
- STAT6 (ang. *signal transducer and activator of transcription 6, interleukin-4 induced*) transduktor sygnału i aktywator 6 transkrypcji indukowany interleukiną 4
- TAC1 (ang. *transmembrane activator and calcium modulator and cyclophilin ligand interactor*) transbłonowy aktywator i modulator wapnia oddziałujący z ligandem dla cyklofiliny
- TadA (ang. *tRNA-specific adenosine deaminase A*) deaminaza A adenozyne specyficzna wobec tRNA
- TDG (ang. *thymidine-DNA glycosylase*) glikozylaza tymidyno-DNA
- TET (ang. *Ten-eleven-translocation methylcytosine dioxygenase*) dioksygenaza 5-metylocytydyny
- TGF- β (ang. *transforming growth factor beta*) transformujący czynnik wzrostu beta
- TLR (ang. *toll-like receptor*) receptor toll-podobny
- TLS (ang. *translesion synthesis polymerase*) polimeraza przeprowadzająca syntezę ponad uszkodzeniami
- TNF α (ang. *tumor necrosis factor alpha*) czynnik alfa nekrozy nowotworów
- TP53BP1(ang. *tumor protein p53 binding protein 1*) białko 1 wiążące białko TP53
- Trim28 (ang. *tripartite motif containing 28*) białko 28 zawierające motyw trójdzielny
- tRNA (ang. *transfer RNA*) transportujący (transferowy) RNA
- UDG (ang. *uracil-DNA glycosylase*) glikozylaza uracylo-DNA
- VIF (ang. *viral infectivity factor*) wirusowy czynnik infekcyjności
- VLRS (ang. *variable lymphocyte receptors*) zmienne receptory limfocytów T i B
- Xpo1 (ang. *exportin 1*) eksportyna 1

1. WPROWADZENIE

Podstawowym wyzwaniem przed jakim stoją wszystkie znane nam formy życia jest stworzenie warunków umożliwiających zachowanie własnego materiału genetycznego w niezmienionej postaci co najmniej do momentu przekazania go następnemu pokoleniu. Nic zatem dziwnego, że w toku ewolucji organizmy rozwinęły szereg skomplikowanych mechanizmów naprawy mutacji, aby minimalizować ilość uszkodzeń i błędów powstających w genomie. Tymczasem, wbrew przedstawionej powyżej regule u kręgowców pojawiła się rodzina białek, mających unikatową zdolność wprowadzania mutacji zarówno do łańcuchów DNA jak i RNA przez deaminację cytydyny do urydyny (rodzina białek AID/APOBEC). Białka AID/APOBEC biorą zatem udział w procesach fizjologicznych, w których wzrost zmienności genetycznej jest korzystny dla organizmu. Ze zjawiskiem tego typu mamy do czynienia w przypadku swoistej oraz nieswoistej odpowiedzi immunologicznej, procesów epigenetycznych lub edycji mRNA. Obecnie najintensywniej badanym przedstawicielem tej stosunkowo niedawno odkrytej rodziny białek jest ludzka deaminaza cytydyny indukowana aktywacją limfocytów B (ang. *activation-induced cytidine deaminase*, AID). Enzym ten, w aktywowanych antygenem limfocytach B, indukuje procesy zmienności przeciwciał prowadząc w konsekwencji do produkcji immunoglobulin specyficznych wobec antygeny i wykazujących różne właściwości efektorowe. Najnowsze dane literaturowe wskazują, że AID poprzez zdolność deaminacji 5mC, może również brać udział w aktywnej globalnej demetylacji genomu. W tym kontekście AID może mieć bezpośredni wpływ na proces reprogramowania epigenetycznego, regulację ekspresji genów i zjawisko pluripotencji komórek. Konsekwencje działania AID, pomimo prostych właściwości katalitycznych tego białka, mogą być zatem zaskakująco istotne.

Kolejną przyczyną, dla której AID przyciąga uwagę wielu naukowców jest potencjał onkogenny wpisany w fizjologiczną funkcję tego białka. AID może bowiem destabilizować genom, wprowadzając do niego liczne mutacje punktowe lub indukując proces jego rearanzacji. Białko to jest więc obiecującym celem terapii nie tylko schorzeń immunologicznych ale i nowotworowych. Mimo że od odkrycia AID upłynęło ponad 15 lat, nadal nie znamy odpowiedzi na wiele podstawowych pytań dotyczących mechanizmów molekularnych leżących u podłoża jej aktywności. Wyniki

badania przeprowadzonych w ramach niniejszej pracy doktorskiej stanowią próbę znalezienia odpowiedzi na niektóre z tych niezwykle fascynujących pytań.

1.1. Rodzina białek AID/APOBEC

U człowieka zidentyfikowano 11 genów kodujących białka należące do rodziny AID/APOBEC: APOBEC1¹ (ang. *apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 1*), APOBEC2, AID (ang. *activation-induced cytidine deaminase*), 7 białek grupy APOBEC3 (od A do H) oraz APOBEC4. Mimo że białka te katalizują tę samą reakcję (deaminacji cytydyny do urydyny), to pełnią różne funkcje biologiczne. AID i APOBEC3 deaminują cytydynę w jednoniciowym DNA (ang. *single-stranded DNA*, ssDNA) biorąc udział w odpowiedzi immunologicznej odpowiednio nabytej i wrodzonej. APOBEC1 edytuje mRNA apolipoproteiny B (białka zaangażowanego w transport lipidów). Funkcje białek APOBEC2 i APOBEC4 nie są znane. Generalnie, nasza wiedza na temat roli biologicznej całej rodziny AID/APOBEC jest bardzo ograniczona.

1.1.1. Ewolucja i organizacja genów

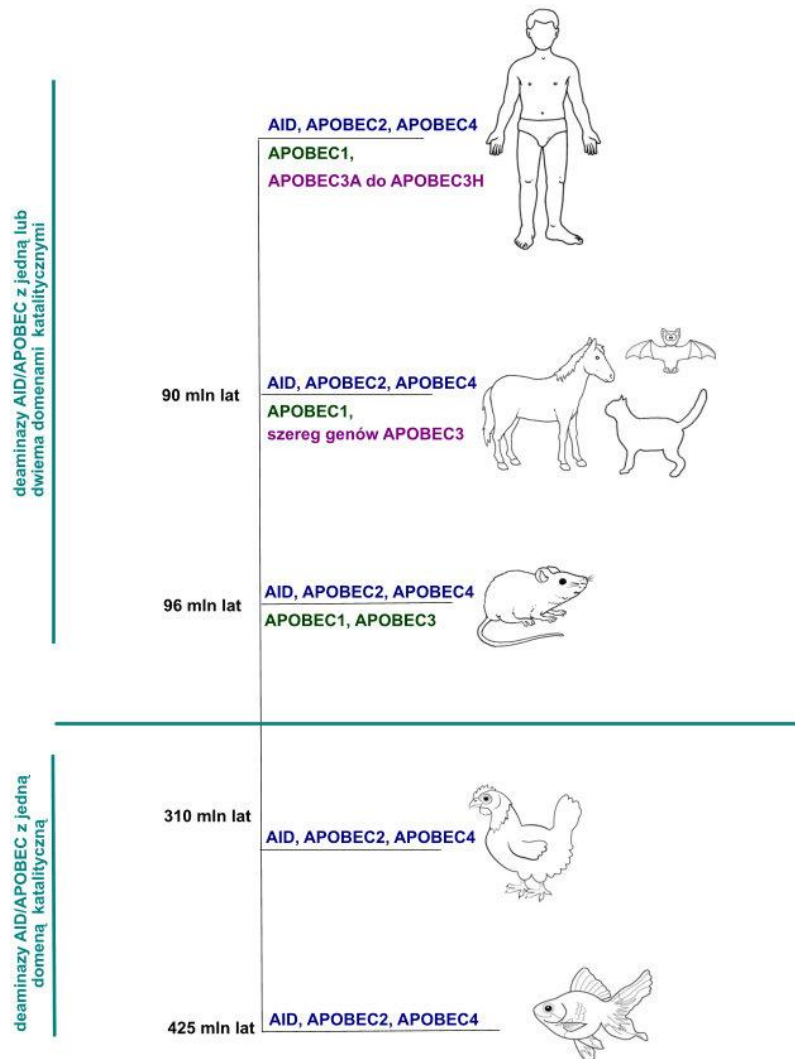
Pierwsze testy aktywności enzymatycznej AID sugerowały, że *in vitro* enzym ten może deaminować wolną cytydynę [1]. Kolejne badania nie potwierdziły jednak tej obserwacji. Obecnie uważa się, że białka z rodziny AID/APOBEC deaminują cytydynę obecną w łańcuchach DNA lub RNA, a wolna cytydyna nie jest ich substratem [2, 3]. Każde białko z tej rodziny ma jedną lub dwie domeny deaminazowe (ang. *catalytic domain*, CD). Domenę katalityczną deaminaz cytydyny tworzy wspólny dla wszystkich białek z tej rodziny motyw: His-X-Glu-X₋₂₃₋₂₈-Pro-Cys-X-Cys, w którym histydyna i dwie cysteiny koordynują atom cynku [2]. Przodek tych białek wyodrębnił się z nadrodziny deaminaz zależnych od cynku już u wczesnych kręgowców [4]. Na podstawie licznych podobieństw strukturalnych i funkcjonalnych, przyjęto uważać, że rodzina białek AID/APOBEC wywodzi się z grupy deaminaz adenozyne (TadA/ADAT2), przekształcających adenozyne do inozyny w tRNA zarówno u *Prokaryota* jak i *Eukaryota*. Białka TadA/ADAT2 modyfikują występującą w antykodonie adenozyne trzydziestą czwartą do inozyny. W efekcie, edytowany tRNA może rozpoznawać więcej niż jeden kodon. O ewolucyjnym pokrewieństwie białek

¹ Rozwinięcie wszystkich skrótów użytych w niniejszej rozprawie doktorskiej znajduje się w wykazie skrótów. Rozwinięcie ważniejsze skrótów zostało ujęte również w tekście.

AID/APOBEC oraz deaminaz adenozytowych działających na łańcuchy polinukleotydowe świadczy kilka faktów. Po pierwsze struktura krystaliczna APOBEC2 wykazuje znaczne podobieństwo do bakteryjnego białka TadA [5, 6]. Sugeruje to zachowawczość mechanizmu, zgodnie z którym rozpoznawane i deaminowane są zasady w łańcuchach DNA/RNA. Po drugie deaminaza adenozyminy (ADAT2) z *Trypanosoma*, w systemie bakteryjnym, wykazuje zdolność deaminacji cytydyny w DNA [7]. Po trzecie powstanie białek AID/APOBEC z deaminaz puryn wydaje się być również bardziej prawdopodobne niż z rodziny deaminaz wolnej cytydyny. Przyjmuje się bowiem, że ewolucyjnie łatwiejszym procesem jest modyfikacja kieszeni katalitycznej niż utworzenie *de novo* struktury rozpoznającej i wiążącej łańcuch polinukleotydowy [2].

Jak już wspomniano, rodzina białek AID/APOBEC pojawiła się po raz pierwszy u wczesnych kręgowców (Rysunek 1.1.1). Ze względu na pełnione funkcje AID i APOBEC3 w mechanizmach odpornościowych, ewolucję tych białek wiąże się z ewolucją układu immunologicznego. AID jest uważana za najstarszego członka rodziny AID/APOBEC. Białko to występuje u ryb kostnoszkieletowych np. *Danio rerio*, a jego homolog pojawia się u bezzuchwoców, których przedstawicielem jest minóg. Sugeruje to, że zaangażowanie białek AID/APOBEC w procesy immunologiczne jest ewolucyjnie starsze niż np. geny immunoglobulin. Przypuszcza się, że u wczesnych kręgowców AID bierze udział w generowaniu zmienności receptorów limfocytarnych (ang. *variable lymphocyte receptors*, VLRs), na których to oparty jest układ odpornościowy tych organizmów [8]. Sekwencje kodujące receptory VLRs zawierają liczne powtórzenia, które ulegają przynajmniej jednemu etapowi różnicowania. Istnieje szereg przesłanek świadczących, że homolog AID może być zaangażowany w ten proces. Wskazują one na ścisłą koewolucję rodziny białek AID/APOBEC i układu immunologicznego kręgowców [4]. Geny kodujące białka APOBEC2 i 4 pojawiają się u ryb chrzęstnoszkieletowych. Ze względu jednak na nieznaną funkcję tych białek, biologiczne znaczenie tego faktu pozostaje zagadką. Znacznie młodsze ewolucyjnie geny białek APOBEC1 i APOBEC3 pojawiają się, odpowiednio, u ssaków i ssaków łożyskowych. APOBEC1 jako jedyny przedstawiciel rodziny AID/APOBEC ma zdolność do edycji RNA *in vivo*. Ponadto białko to zachowuje częściową zdolność do edycji ssDNA [9]. Aktywność ta może wskazywać na dodatkową, jeszcze nieodkrytą funkcję APOBEC1. Najmłodszą podrodzinę tworzą białka APOBEC3. Ich ewolucja u ssaków łożyskowych jest unikalnym przykładem

ukazującym w jaki sposób zduplikowane geny mogą nabywać nowych funkcji pod wpływem presji selekcyjnej.



Rysunek 1.1.1. Pojawianie się kolejnych genów białek AID/APOBEC w trakcie filogenezy kręgowców. Na podstawie [4, 10]

Pelniejszy wgląd w ewolucję rodziny białek AID/APOBEC daje analiza struktury kodujących je genów. Geny większości białek z rodziny AID/APOBEC składają się z pięciu eksonów i jest to cecha wspólna z grupą genów deaminaz TadA/ADAT2 oraz genów deaminaz wolnej cytydyny. Miejsce katalityczne tych enzymów kodowane jest przez ekson trzeci. W przeciwieństwie do pozostałych białek rodziny AID/APOBEC, geny kodujące APOBEC2 i 4 mają odpowiednio tylko dwa i trzy eksony, przy czym uważa się, że sekwencja kodująca mieści się głównie w obrębie eksonu drugiego (kilka reszt aminokwasowych kodowanych przez ekson

pierwszy *APOBEC2*² nie wykazuje podobieństwa do żadnej znanej sekwencji aminokwasowej). Ograniczenie liczby eksonów w *APOBEC2* i *4* wskazuje, że geny te powstały na drodze wczesnej retrotranspozycji. *APOBEC1* i *3*, które wyewoluowały później, powstały na drodze kilku następujących po sobie duplikacji *locus AID*. *Locus APOBEC1* powstało przez odwrotną duplikację *locus AID* na tym samym chromosomie (u człowieka chromosom 12). U większości ssaków oba *loci* dzieli region o długości około 40 kpz, a u naczelnych o długości około 1 Mpz. *APOBEC1* w tej samej co *AID* orientacji występuje u torbaczy, *locus APOBEC3* pojawia się podczas dywergencji linii torbaczy i łozyskowców. Pierwotna duplikacja genu zajmującego to *locus* u przodka ssaków łozyskowych uformowała dwa geny: *APOBEC3* i *3s*, które następnie ewoluowały przez wielokrotne duplikacje i fuzje. U niektórych grup zwierząt, takich jak: gryzonie, świniowate i krętorogie, te dwa pierwotne geny połączyły się tworząc jeden, z dwiema domenami katalitycznymi. U innych grup, takich jak: naczelne, koniowate, nietoperze i kotowate, jeden z dwóch genów uległ wielokrotnej duplikacji tworząc szereg genów *APOBEC3*. U naczelnych *locus APOBEC3* jest największe. Składa się na nie 7 obecnie znanych genów białek *APOBEC3*. Przypuszcza się, że główny wpływ na ewolucję *APOBEC3* u naczelnych miała presja selekcyjna wywierana przez retrowirusy i retroelementy. Genomy retrowirusowe i retroelementy są bowiem obiektami docelowymi dla białek *APOBEC3* [2, 4, 11-13].

1.1.2. Funkcje

Pierwszym odkrytym białkiem rodziny *AID/APOBEC* było białko *APOBEC1*. Białko to potranskrypcyjnie modyfikuje mRNA apolipoproteiny B (apoB) przez deaminację cytydyny w pozycji 6666. *APOBEC1* generuje w ten sposób dodatkowy kodon STOP. W rezultacie powstają dwie alternatywne formy apoB. Każda z form apoB pełni inną funkcję biologiczną: dłuższa apoB-100 jest odpowiedzialna za transport endogennie syntetyzowanych triglicerydów i cholesterolu, a skrócona apoB-48 bierze udział we wchłanianiu i transporcie egzogennych triglicerydów oraz wchodzi w skład chylomikronów. Choć mechanizm edycji mRNA przez *APOBEC1* został pokazany jedynie dla mRNA apoB, to hipotetycznie białko to może brać udział również w potranskrypcyjnym edytowaniu innych mRNA [14].

² W niniejszej rozprawie doktorskiej przyjęto zasadę pisowni nazw genów kursywą, a nazw białek bez kursywy. Dla zachowania czytelności pracy pominięto zasadę różnicowania pisowni genów ludzkich i zwierzęcych przy pomocy wielkości liter.

Funkcja biologiczna APOBEC2 i APOBEC4 nie jest znana. *APOBEC2*, znany również jako *ARCD-1*, ulega ekspresji głównie w mięśniu sercowym i mięśniach szkieletowych. Jak dotąd substrat *in vivo* dla APOBEC2 nie został zidentyfikowany. Podczas gdy inne białka z tej rodziny przejawiają aktywność deaminazową nawet przy braku kofaktorów, APOBEC2 nie wykazuje aktywności enzymatycznej w warunkach *in vitro*. Obecność APOBEC2 jest jednak kluczowa dla rozwoju mięśni co wskazuje na jego nieokreśloną ale specyficzną funkcję w mięśniu sercowym i mięśniach szkieletowych [15]. Pokazano również, że transgeniczne myszy z nadprodukcją tego białka rozwijają nowotwory płuc i wątroby [16]. O istnieniu genu *APOBEC4* świadczą jedynie analizy bioinformatyczne. Jak dotąd jego ekspresji nie stwierdzono w żadnej tkance. Sugeruje się jedynie możliwość jego ekspresji w jądrach. Podobnie jak APOBEC2, białko APOBEC4 nie przejawia aktywności *in vitro*. Być może jest ono funkcjonalne tylko w obecności specyficznych kofaktorów [15, 17, 18].



Rysunek 1.1.2. Organizacja domenowa deaminazy cytydyny indukowanej aktywacją limfocytów B (AID). Organizacja domenowa została przewidziana na podstawie podobieństwa sekwencji z innymi białkami AID/APOBEC. Na rysunku zaznaczono poszczególne domeny z wyszczególnieniem ich pozycji w sekwencji aminokwasowej AID. NLS – sygnał lokalizacji jądrowej, NES – sygnał eksportu jądrowego, H56-C87-C90 – pozycje motywu koordynującego atom cynku w miejscu aktywnym. Na podstawie [19].

Deaminaza cytydyny indukowana aktywacją limfocytów B (ang. *activation-induced cytidine deaminase*, AID) jest najlepiej poznany przedstawicielem rodziny AID/APOBEC. Jest to białko o masie 24 kDa, którego struktura trzeciorzędowa nie została jak dotąd poznana. Rysunek 1.1.2. ilustruje organizację domen AID, przewidywaną na podstawie podobieństwa sekwencji z innymi białkami AID/APOBEC. W świetle dotychczasowych badań wydaje się, że podstawową funkcją AID jest inicjowanie procesów odpowiedzialnych za zmienność przeciwciał. Mechanizmy molekularne leżące u podłoża tej aktywności zostaną szczegółowo omówione w dalszych rozdziałach tej pracy.

Tabela 1.1.1. Białka rodziny AID/APOBEC. Porównanie ich funkcji, preferowanych motywów sekwencyjnych w obrębie cząsteczek DNA/RNA oraz sekwencji kodujących pętlę odpowiedzialną za specyficzność substratową (opis w podrozdziale 1.1.3). Na podstawie [15, 17].

BIĄŁKO	PREFEROWANY MOTYW SEKWENCYJNY	SEKWENCJA AMINOKWASOWA PĘTLI ODPOWIEDZIALNEJ ZA SPECYFICZNOŚĆ SUBSTRATOWĄ	FUNKCJA
AID	WRCY* (DNA)	¹¹² RLYFCEDRK	inicjowanie procesów zmienności przeciwciał; możliwa inhibicja retroelementów i wirusów takich jak: HSV ³ , HBV
APOBEC1	C6666 (ApoB mRNA) G/CTC (DNA)	²⁴⁰ RLFWHMDQQ	potranskrypcyjna modyfikacja mRNA apoB; możliwa inhibicja retroelementów i wirusów takich jak: HSV, HBV
APOBEC2	Nieznany	¹⁵³ RLFMWEEPE	nieznana
APOBEC3A	T/CCA (DNA)	¹²³ RIYDYDPL	inhibicja retroelementów i wirusów takich jak: HSV, HBV, HPV, HIV, MLV, AAV
APOBEC3B	C/GTC (DNA)	¹²² RLYYYWVERD (CD1) ³¹¹ RIYDYDPL (CD2)	inhibicja retroelementów i wirusów takich jak: HBV
APOBEC3C	TC/TC (DNA)	¹²² RLYYFQYPC	inhibicja retroelementów i wirusów takich jak: HSV, HBV, HPV
APOBEC3DE	WWC (DNA)	¹³⁵ RLYYYRDR (CD1) ³¹⁹ RLCYFWDTD (CD2)	inhibicja wirusów takich jak: HIV, SIV
APOBEC3F	TTC (DNA)	¹²¹ RLYYYWVERD (CD1) ³⁰⁵ RLYYFWDTD (CD2)	inhibicja retroelementów i wirusów takich jak: HBV, HIV, SIV, EIAV, FV
APOBEC3G	CCC (DNA)	¹²² RLYYFWDPD (CD1) ³¹³ RIYDDQGR (CD2)	inhibicja retroelementów i wirusów takich jak: HSV, HBV, HIV, SIV, MLV, EIAV, FV
APOBEC3H	Nieznany	¹¹¹ RLYYHWCKP	inhibicja wirusów takich jak: HBV, HPV, HIV, SIV
APOBEC4	Nieznany	brak podobieństwa	nieznana

* W = A/T; R = A/G; Y = C/T

Grupa białek APOBEC3 jest zaangażowana w unieczynnianie egzogen- i endogennych zagrożeń, związanych z aktywnością retrowirusów i elementów ruchomych w genomie. Tym samym białka te pełnią zasadnicze funkcje

³ Rozwinięcie skrótów nazw wirusów znajduje się w wykazie skrótów.

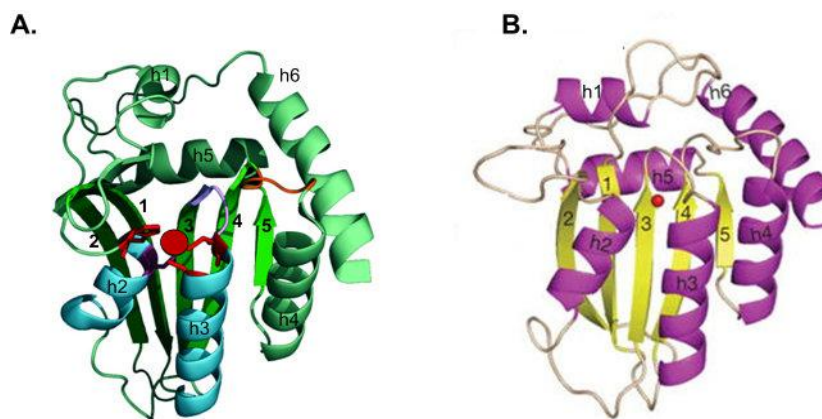
w mechanizmach odporności wrodzonej [20]. Pierwszym dowodem przemawiającym za taką właśnie rolą białek APOBEC3 było odkrycie, że APOBEC3G bierze udział w unieczynnieniu wirusa HIV pod nieobecność wirusowego białka VIF. Białko APOBEC3G deaminuje wirusowy DNA i w ten sposób zaburza syntezę formy dwuniciowej lub prowadzi do jej degradacji przez komórkowe enzymy naprawcze. Co ciekawe, w odpowiedzi wirus HIV wytworzył mechanizm przeciwdziałający anty-retrowirusowej aktywności APOBEC3G. Białko VIF wiąże APOBEC3G i indukuje jego degradację w proteasomie oraz zapobiega jego wbudowywaniu do wirionów potomnych⁴ [21, 22]. Ludzkie APOBEC3G nie jest zatem skuteczne wobec wirusa HIV. W badaniach *in vitro* skutecznie unieczynnia ono małpiego wirusa niedoboru odporności (SIV). Przypuszcza się, że polimorfizm tego białka jest źródłem bariery międzygatunkowej przy przenoszeniu się SIV u naczelnych [10, 23, 24]. Wszystkie białka APOBEC3 stanowią element mechanizmów odpowiedzialnych za odporność wrodzoną. Enzymy te deaminują genomy wirusów, różniąc się przy tym preferowanym kontekstem sekwencyjnym, w którym usytuowana jest podlegająca modyfikacji cytydina. Hamują one również przenoszenie się retroelementów w genomie. Zestawienie funkcji i preferowanych motywów sekwencyjnych poszczególnych białek APOBEC3 zawiera Tabela 1.1.1. [20]. W niektórych układach doświadczalnych również APOBEC1 i AID wykazują aktywność względem retroelementów i wirusów [25-29]. Sugeruje to, że i te białka mogą warunkować odporność wrodzoną.

1.1.3. Struktura i właściwości biochemiczne

Wszystkie białka z rodziny AID/APOBEC katalizują tę samą reakcję, choć pełnią odmienne funkcje biologiczne. Jak już wspomniano, centrum katalityczne tych enzymów tworzy charakterystyczny motyw: His-X-Glu-X₋₂₃₋₂₈-Pro-Cys-X-Cys, w obrębie którego histydyna i dwie cysteiny koordynują atom cynku [2]. Postulowany mechanizm katalizowanej przez AID/APOBEC reakcji deaminacji cytydyny do urydyny zakłada nukleofilowy atak hydroksylu na grupę aminową przy węglu C4. Hydroksyl ten powstaje w wyniku aktywacji i rozczepienia cząsteczki wody (koordynowanej przez kation cynkowy (Zn²⁺)) przy udziale zdeprotonowanej grupy karboksylowej reszty glutaminianu E58. E58 działa następnie jako donator protonu

⁴ Pod nieobecność wirusowego białka VIF, APOBEC3G wiąże się do wirusowego RNA i wraz z nim zostaje wbudowane do wirionów potomnych. W kolejnej zakażonej komórce białko to zaburza odwrotną transkrypcję deaminując powstające DNA.

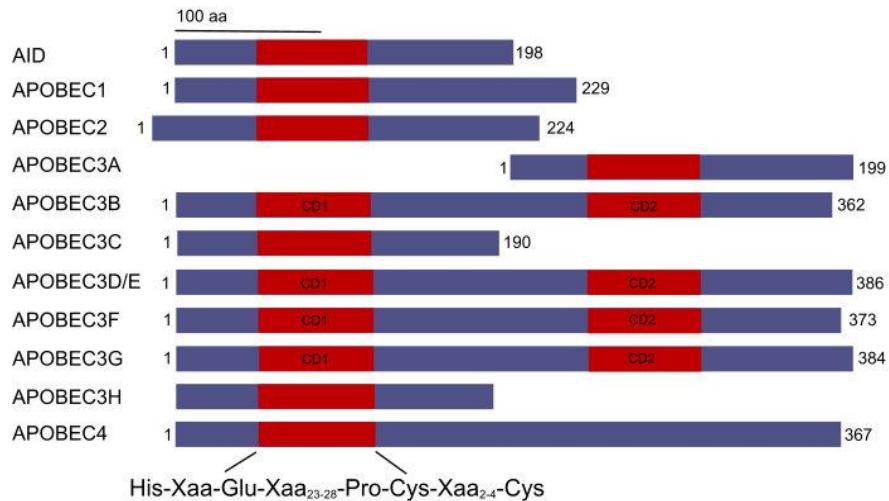
wobec azotu N3 tworząc z nim wiązanie wodorowe. Hydroksyl natomiast jest donorem protonu dla odłączanej grupy aminowej [3]. Mechanizm ten został zaproponowany na podstawie informacji uzyskanych dla bakteryjnej deaminazy działającej na wolną cytydynę. Wydaje się on być wspólny dla całej nadrodziny deaminaz cytydyny zależnych od cynku.



Rysunek 1.1.3. Struktury krystaliczne białek z rodziny AID/APOBEC. A) Struktura krystaliczna niescharakteryzowanego funkcjonalnie APOBEC2 (PDB 2nyt) [5]. B) Struktura krystaliczna aktywnej domeny katalitycznej białka APOBEC3G (PDB 3e1u) [30]. Na rysunkach cyframi od 1 do 5 zaznaczono kolejne harmonijki β . Helisy α oznaczono h1-h6. Czerwone punkty uwidaczniają położenie kationów cynkowych. Reprodukacja za zgodą © Nature Publishing Group [5, 30].

Biorąc pod uwagę poznane struktury krystaliczne białek AID/APOBEC, czyli strukturę niescharakteryzowanego funkcjonalnie APOBEC2 (Rysunek 1.1.3.A) oraz strukturę aktywnej domeny katalitycznej białka APOBEC3G (Rysunek 1.1.3.B), przypuszcza się, że kieszeń katalityczną białek z tej rodziny tworzy charakterystyczny układ struktur drugorzędowych – pięć harmonijek β i dwie helisy α , w obrębie których umiejscowiona jest histydyna i dwie cysteiny koordynujące atom cynku [31]. Interesujące jest również, że niektóre białka APOBEC3 (dokładnie APOBEC3B, 3DE, 3F, 3G) posiadają dwie domeny katalityczne, przy czym N-terminalna domena katalityczna (CD1) APOBEC3G i 3F jest nieaktywna. Umiejscowienie domen katalitycznych w obrębie poszczególnych białek AID/APOBEC ukazano schematycznie na Rysunku 1.1.4. Nieaktywne katalitycznie domeny białek APOBEC3 posiadają w swej strukturze wszystkie reszty aminokwasowe niezbędne do aktywności katalitycznej. Nie znamy zatem strukturalnych podstaw braku ich aktywności. W przypadku APOBEC3G nieaktywna domena CD1 bierze udział w wiązaniu wirusowego RNA i inkorporacji białka do wirionu, podczas gdy aktywna domena CD2

deaminuje ssDNA⁵ [31]. Można zatem założyć, że nieaktywna domena w białkach rodziny AID/APOBEC wiąże RNA, aktywna natomiast deaminuje cytydynę w obrębie ssDNA.



Rysunek 1.1.4. Schemat obrazujący położenie domeny katalitycznej w obrębie poszczególnych białek AID/APOBEC. Czerwone pole reprezentuje domenę katalityczną (CD). Zaznaczono również liczbę reszt aminokwasowych tworzących poszczególne białka.

Pojedyncze domeny deaminaz AID/APOBEC wiążą ssDNA jak i ssRNA, przy czym zdolność deaminacji RNA wykazano jedynie dla białka APOBEC1 i szczurzego APOBEC3 [4]. Pomimo poznania struktury trzeciorzędowej fragmentu białka APOBEC3G zawierającego aktywną domenę CD2 ciągle nie jest jasne, dlaczego białko to nie ma zdolności edycji RNA. Istnieją natomiast przesłanki, że RNA zaburza aktywność deaminazową AID i APOBEC3G [32, 33]. Mechanizm hamowania aktywności deaminazowej AID i APOBEC3 przez RNA, może być swoistym sposobem regulacji tej aktywności. Wiązanie RNA przez te białka, nasuwa jednak pytania o możliwość jego edycji w bardziej skomplikowanych kompleksach białkowych (jak ma to miejsce w przypadku APOBEC1⁶). Po nieudanych próbach wykazania, że białka APOBEC3 i AID są zdolne do deaminacji RNA w systemach *in vitro*, zaniechano dalszych tego typu badań [32].

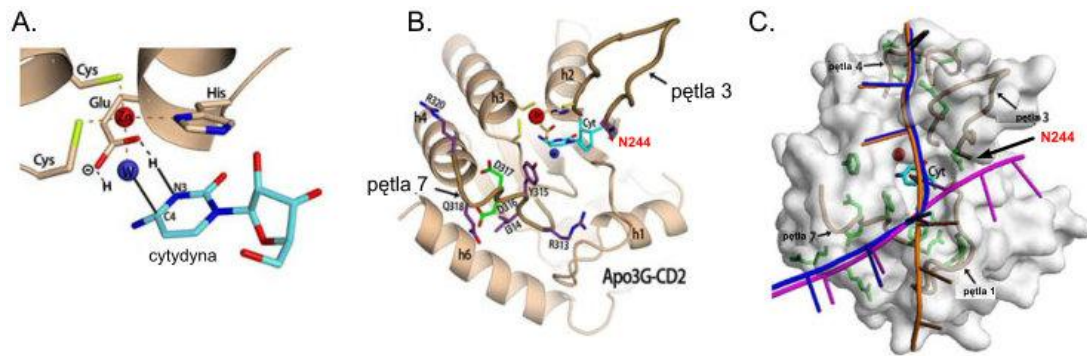
Pomimo iż mechanizm katalizowanej reakcji wydaje się być wspólny dla całej rodziny AID/APOBEC, to każde z białek wykazuje preferencję wobec charakterystycznego motywu sekwencyjnego, w obrębie którego ulokowana jest

⁵ Związane z retrowirusowym RNA białko APOBEC3G jest wbudowywane do wirionów potomnych. W kolejnej zakażonej komórce białko to zaburza odwrotną transkrypcję deaminując powstające DNA.

⁶ APOBEC1 wraz z innymi białkami tworzy kompleks zwany edytosomem 27S (patrz dalsza część tego podrozdziału).

podlegająca deaminacji cytydyna (Tabela 1.1.1) [31]. Czynniki determinujące specyficzność tych białek wobec poszczególnych motywów sekwencyjnych nie zostały jak dotąd w pełni poznane. Przypuszcza się, że za ich rozpoznawanie odpowiedzialne są reszty aminokwasowe tworzące pętlę 7 (pętlę pomiędzy harmonijką $\beta 4$ i helisą $\alpha 4$). Wykazano bowiem, że wymiana tej pętli w białku AID na odpowiadającą jej pętlę z aktywnej domeny katalitycznej CD2 APOBEC3G lub 3F, zmienia spektrum motywów sekwencyjnych preferowanych przez AID [34]. Zaangażowanie tych reszt aminokwasowych w zapewnienie specyficzności substratowej pokazano również wykorzystując szereg mutantów białek APOBEC3 [31]. Zakłada się, że 9-11 reszt aminokwasowych tworzących pętlę specyficzności substratowej wiąże nukleotydy od strony 5' cytydyny podlegającej deaminacji i w ten sposób pomaga ją pozycjonować w centrum katalitycznym enzymu. Reszty aminokwasowe pętli 7 mogą oddziaływać z łańcuchem nukleotydowym zarówno na zasadzie oddziaływań hydrofobowych, jonowych jak i warstwowych [17]. Porównanie sekwencji aminokwasowych tworzących pętlę 7 dla poszczególnych białek AID/APOBEC zawiera Tabela 1.1.1, a lokalizację tej pętli w strukturze CD2 APOBEC3G zaznaczono na Rysunku 1.1.5.B.

Aby zaszła reakcja deaminacji, cytydyna musi wniknąć głęboko do kieszeni katalitycznej enzymu. Położenie cytydyny w miejscu aktywnym CD2 APOBEC3G przedstawia Rysunek 1.1.5.A. Pozycja cytydyny została zaproponowana na podstawie nałożenia struktur krystalicznych CD2 APOBEC3G (PDB 3e1u) oraz kompleksu mysiej deaminazy wolnej cytydyny (ang. *cytidine deaminase*, CDA) z substratem (PDB 2fr6) [30, 31]. Na Rysunku 1.1.5.B przedstawiono aranżację drugorzędowych elementów strukturalnych otaczających miejsce aktywne CD2 APOBEC3G z zaznaczeniem reszt aminokwasowych tworzących pętlę specyficzności substratowej. Na szczególną uwagę zasługuje reszta N244 w obrębie pętli 3 CD2 APOBEC3G. Jest ona wysoce zachowawcza w całej rodzinie deaminaz zależnych od cynku i odpowiada reszcie N51 w strukturze AID. Według danych literaturowych mutacja reszty N244 białka APOBEC3G lub też N51 białka AID do alaniny całkowicie znosi aktywność deaminazową *in vitro* obu enzymów [30, 31, 35]. Pokazano również, że reszta N244 bezpośrednio oddziałuje z podlegającą deaminacji cytydyną [31, 36].



Rysunek 1.1.5. Wiązanie substratu przez białka z rodziny AID/APOBEC. A) Zaproponowany model położenia cytydyny w miejscu aktywnym białek AID/APOBEC w trakcie reakcji deaminacji. Model powstał przez nałożenie struktur krystalicznych CD2 APOBEC3G (PDB 3e1u) oraz kompleksu mysiej deaminazy wolnej cytydyny z substratem (PDB 2fr6). Czerwony punkt na rysunku reprezentuje atom cynku; niebieski punkt – cząsteczkę wody. B) Struktura krystaliczna aktywnej domeny katalitycznej białka APOBEC3G z zaznaczeniem: struktur drugorzędowych otaczających miejsce aktywne, reszt aminokwasowych w obrębie pętli specyficzności substratowej (pętli 7) oraz reszty N244 w obrębie pętli 3 (opis w tekście). Kolorem niebieskim na rysunku zaznaczona jest cytydyna podlegająca deaminacji. C) Zaproponowane dla CD2 APOBEC3G modele wiązania ssDNA. Kolorem różowym oznaczono ssDNA związany z CD2 w sposób zaproponowany przez Holdena [30]. Analogicznie kolorem pomarańczowym oznaczono ssDNA związany z CD2 w sposób zaproponowany przez Chena i Furokawę [37, 38]. ssDNA związany z CD2 w sposób wynikający z obu hipotez oznaczono kolorem granatowym [31]. Reprodukacja za zgodą © Springer Science and Business Media [31].

Na podstawie struktury krystalicznej CD2 APOBEC3G lub struktury tej samej domeny określonej metodą magnetycznego rezonansu jądrowego (ang. *nuclear magnetic resonance*, NMR) zaproponowane zostały różne sposoby wiązania ssDNA przez to białko (Rysunek 1.1.5.C.). Z sugerowanych modeli, model Holdena (stworzony na podstawie badań krystalograficznych) wydaje się być najbardziej zasadny [30]. W modelu Chena i Furukawy (opracowanym na podstawie analiz NMR) ssDNA nie przechodzi bowiem w okolicy pętli specyficzności substratowej [37, 38]. Możliwa jest również kombinacja obu modeli [31]. Niestety nasza dotychczasowa wiedza nie jest wystarczająca do jednoznacznego rozstrzygnięcia tego problemu. Poznanie struktury kompleksu białka z rodziny AID/APOBEC z substratem pozwoliłoby zmienić tę sytuację.

Kolejną niewiadomą pozostaje mechanizm skanowania nici DNA/RNA przez deaminazy AID/APOBEC. Część badaczy postuluje możliwą procesywność i kierunkowość ich działania. Procesywność w przypadku APOBEC3G i AID jest definiowana jako zdolność do deaminacji wielu cytydyn w obrębie ssDNA. Dopiero po

ich modyfikacji enzym oddysocjowuje by związać się z innym substratem [39]. Taki sposób działania AID i APOBEC3G jest postulowany od wielu lat przez grupę Goodmana [39, 40]. Ta sama grupa zaproponowała mechanizm skanowania ssDNA przez AID poprzez krótkie ślizgi i przeskoki. Dla APOBEC3G badacze postulują większy udział dłuższych przeskoków lub też specyficzny mechanizm transferu pomiędzy dwoma segmentami ssDNA (ang. *intersegmental transfer mechanism*) zaproponowany przez Nowarskiego w 2008 roku [41]. AID działa w obrębie niewielkiego bąbla transkrypcyjnego, a APOBEC3G na stosunkowo długiej nici wirusowego DNA. Nie dziwią zatem pewne potencjalne różnice w działaniu obu tych białek. Ponadto, dla APOBEC3G zaobserwowano, że białko to w sposób losowy wiąże się do cząsteczki substratowej, ale częstość wprowadzanych przez nie mutacji rośnie w kierunku końca 5'. Na tej podstawie badacze określili, że aktywność deaminazową APOBEC3G cechuje kierunkowość 3'→5' [42]. Nie obserwuje się natomiast takiej kierunkowości dla AID. Przypuszcza się, że za obserwowaną kierunkowość działania APOBEC3G odpowiedzialna jest nieaktywna domena katalityczna CD1. Skrócona forma białka, obejmująca jedynie aktywną domenę katalityczną (CD2) jest pozbawiona tej właściwości [30]. Na tej podstawie można przypuszczać, że wszystkie białka AID/APOBEC z jedną domeną katalityczną są pozbawione kierunkowości działania [17].

Nie ma zgodności wśród badaczy co do dimeryzacji/oligomeryzacji białek z rodziny AID/APOBEC. Wykazano, że APOBEC1 podczas edycji mRNA apolipoproteiny B działa w formie dimeru. Dimer ten oddziałuje z dimerem białek ACF tworząc tzw. edytosom 27S rozpoznający specyficzną 11-nukleotydową sekwencję [15]. Poznana struktura krystaliczna APOBEC2 jest strukturą tetrameru, ale ponieważ białko to nie zostało scharakteryzowane funkcjonalnie, nie ma pewności, że jest to forma aktywna enzymatycznie. Na podstawie struktury krystalicznej CD2 APOBEC3G postuluje się dimeryzację (oddziaływanie typu „głowa do głowy” poprzez domeny CD1) białka w sposób zależny od RNA. Nie jest również jasne czy AID funkcjonuje jako monomer, dimer czy oligomer. Czwartorzędowa struktura nie wydaje się być jednak konieczna do aktywności enzymatycznej białek AID i APOBEC3 [40, 43, 44]. Poznanie struktury krystalicznej deaminazy AID/APOBEC w kompleksie z substratem pozwoliłoby rozwikłać również i tę kwestię.

1.1.4. Podsumowanie

Pomimo że cała rodzina białek AID/APOBEC jest obiektem intensywnych badań, ciągle niewiele wiadomo na temat ich funkcji, regulacji aktywności oraz wpływie na stabilność i ewolucję genomu. Indukowana przez AID/APOBEC niestabilność DNA z jednej strony stanowi zagrożenie dla organizmu, z drugiej strony jest źródłem korzystnej zmienności (np. przeciwciał) lub też mechanizmem obronnym przed retrowirusami i elementami ruchomymi w genomie. Uderzająca jest prostota reakcji katalizowanej przez białka AID/APOBEC w zestawieniu z plejotropowym efektem ich działania. Obserwacja ta stała się inspiracją do pogłębionych badań będących przedmiotem niniejszej rozprawy. Jako ich obiekt wybrane zostało białko AID będące najstarszym przedstawicielem rodziny.

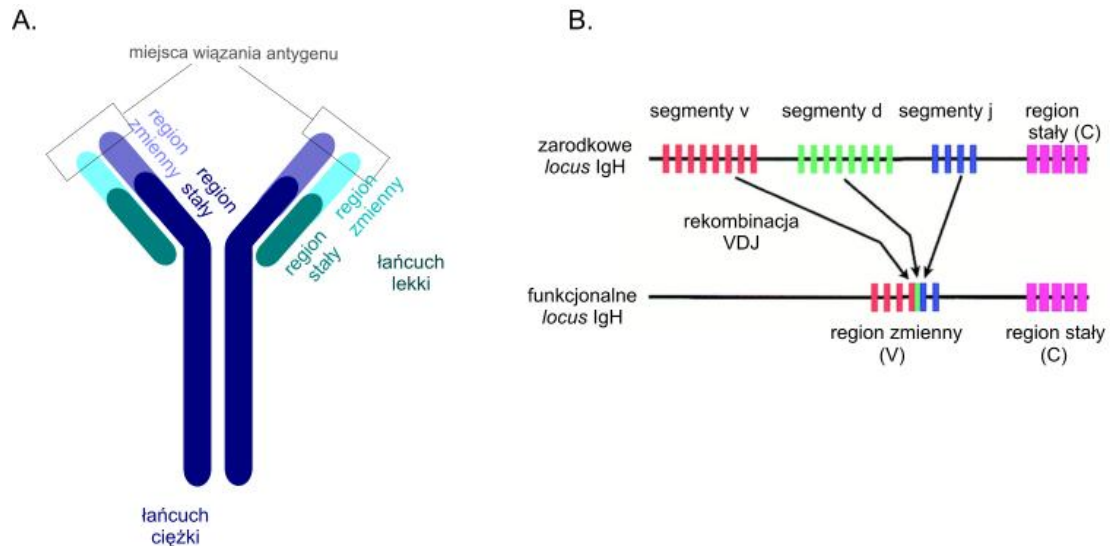
1.2. Odkrycie AID oraz roli jaką białko to odgrywa w procesie powstawania przeciwciał

Efektywna humoralna odpowiedź immunologiczna wymaga zdolności do produkcji niezwykle szerokiego repertuaru przeciwciał rozpoznających patogeny. Tak wielka różnorodność osiągana jest dzięki istnieniu szeregu mechanizmów różnicowania przeciwciał, poprzez zwiększenie poziomu zmienności kodujących je genów.

Można wyróżnić dwa główne etapy procesu różnicowania przeciwciał: pierwszy – niezależny od obecności antygeny i drugi – uwarunkowany obecnością antygeny. W różnicujących się limfocytach B i T, dochodzi do pierwotnej rearanzacji segmentów genowych: v, d, j (tzw. rekombinacji VDJ) (Rysunek 1.2.1). Prowadzi ona do powstania kombinacji segmentów kodujących region zmienny (V) zarówno immunoglobulin jak i receptorów ulokowanych na limfocytach B i T. Rekombinacja ta umożliwia zatem powstanie pierwotnego, niezależnego od obecności antygeny, zestawu przeciwciał i specyficznych dla limfocytów receptorów powierzchniowych. W odróżnieniu od limfocytów T, w aktywowanych przez antygen limfocytach B, geny immunoglobulin ulegają drugiej rundzie różnicowania. Proces ten zachodzi w centrum namnażania limfocytów (zwanym również ośrodkiem odczynowym; ang. *germinal center*, GC) w grudkach limfatycznych. Ta tzw. drugorzędowa zmienność przeciwciał pozwala na produkcję immunoglobulin wysoce specyficznych względem antygeny oraz umożliwia ich specjalizację funkcjonalną (powstawanie przeciwciał różnych klas).

Za proces drugorzędowego różnicowania przeciwciał odpowiada zjawisko: (i) hiperzmienności somatycznej (ang. *somatic hypermutation*, SHM), (ii) rekombinacyjnego przełączania klas (ang. *class switch recombination*, CSR). SHM polega na wprowadzeniu mutacji (zwykle jednonukleotydowych substytucji) w regionie zmiennym (V) genów immunoglobulin. Region zmienny w przeciwciele jest odpowiedzialny za wiązanie antygeny i tylko limfocyty B produkujące przeciwciała o wysokiej specyficzności względem antygeny przeżywają i ulegają proliferacji. W ten sposób proces SHM wraz z selekcją klonalną limfocytów B jest gwarantem wytworzenia swoistej odporności na dany patogen. CSR polega na wymianie fragmentu genu kodującego część stałą łańcucha ciężkiego immunoglobuliny. Ten element przeciwciała określa jego klasę, stąd zjawisko to nazwano przełączaniem klas. Poszczególne klasy immunoglobulin mają odmienne powinowactwo wobec receptorów komórkowych, a co za tym idzie, pełnią odmienne funkcje podczas aktywowania szlaków odpowiedzi immunologicznej. Molekularne podłoże SHM i CSR pozostawało

niepoznane aż do końca XX wieku, kiedy to wyniki badań grupy kierowanej przez Tasuku Honjo doprowadziły do odkrycia deaminazy cytydyny indukowanej aktywacją limfocytów B.



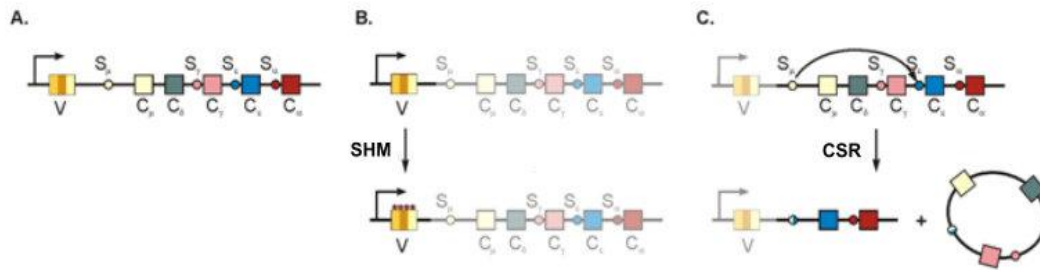
Rysunek 1.2.1. Budowa przeciwciał i organizacja kodujących je genów. A) Schemat budowy przeciwciała. Przeciwciało zbudowane jest z czterech glikozylowanych łańcuchów peptydowych połączonych mostkami dwusiarczkowymi: dwóch dłuższych łańcuchów ciężkich (zaznaczonych na rysunku kolorem granatowym) oraz dwóch krótszych łańcuchów lekkich (zaznaczonych na rysunku kolorem turkusowym). W każdym z powyższych łańcuchów wyróżnia się region stały (ciemniejszy odcień) oraz rozpoznający antygen region zmienny (jaśniejszy odcień); B) Schemat procesu rekombinacji VDJ. Podstawą tego procesu jest istnienie wielu różnych genów (podzielonych na segmenty v, d i j), które w rezultacie łączenia się ze sobą (na drodze rekombinacji) tworzą funkcjonalny gen kodujący region zmienny przeciwciała (V). Region stały przeciwciała (C) nie ulega rekombinacji VDJ. Rekombinacja VDJ zachodzi w sposób analogiczny dla łańcucha ciężkiego i lekkiego z tą różnicą, iż w przypadku łańcucha lekkiego brak jest segmentów d. Na podstawie [45].

Białko AID zostało odkryte w 1999 roku podczas poszukiwań czynnika powstającego specyficznie tylko w aktywowanych przez antygen limfocytach B, w których to indukowane jest zjawisko SHM. [1]. Bazując na homologii AID i APOBEC1, początkowo sądzono, że enzym ten edytuje RNA. Obecnie uważa się, że AID deaminuje cytydynę do urydyny w odcinkach ssDNA powstających podczas rozplatania helisy DNA w procesie transkrypcji [32, 46, 47]. Kolejne doświadczenia pokazały, że AID jest niezbędna do inicjowania dwóch kluczowych procesów różnicowania przeciwciał: SHM oraz CSR. Brak funkcjonalnego białka AID skutkuje zespołem HIGM2 (ang. *hyper IgM syndrome 2*), który objawia się powstawaniem

wyłącznie przeciwciał klasy IgM [48] i związany jest z szeregiem zaburzeń w odpowiedzi immunologicznej.

Obecnie przyjmuje się, że AID inicjuje procesy SHM i CSR przez deaminację cytydyny do urydyny, odpowiednio, w zmiennym (V) lub stałym (C) regionie kodującym immunoglobuliny. AID deaminuje cytydyny preferencyjnie w obrębie motywu sekwencyjnego WRCY (W = A/T; R = A/G; Y = C/T), i to zarówno *in vivo*, gdy uczestniczy w procesach różnicowania przeciwciał, jak i w badaniach *in vitro* [32]. Konsekwencją deaminacji C jest powstanie pierwotnego niesparowania G-U, które może mieć kilka następstw. W połączeniu z replikacją może powodować powstawanie mutacji typu tranzycji (C-G do T-A) (tzw. faza 1a SHM). Urydyna może również zostać rozpoznana i wycięta przez glikozylazę UDG (ang. *uracil-DNA glycosylase*, UDG), która działa w ramach mechanizmu naprawy mutacji przez wycięcie zasad (ang. *base excision repair*, BER) (tzw. faza 1b SHM) [46, 49]. Co ciekawe, w komórkach ssaków występują cztery enzymy zdolne do rozpoznawania i usuwania urydyny w DNA: wspomniana już glikozylaza UDG, jednofunkcyjna glikozylaza uracylo-DNA selektywna wobec ssDNA (ang. *single – strand selective monofunctional uracil-DNA glycosylase*, SMUG), glikozylaza tymidyno-DNA (ang. *thymidine-DNA glycosylase*, TDG) i glikozylaza MBD4 (ang. *methyl-CpG binding domain protein 4*), które potencjalnie mogą komplementować swoje funkcje [50]. UDG wydaje się być jednak główną glikozylazą w procesach różnicowania przeciwciał. Mutacje w genie *UDG* powodują zespół HIGM5, który podobnie jak zespół HIGM2 wiąże się z szeregiem zaburzeń w odpowiedzi immunologicznej i powstawaniem przeciwciał wyłącznie klasy IgM. Działanie wszystkich wymienionych glikozylaz prowadzi do powstania miejsca apirymidynowego, które może stać się substratem dla polimerazy DNA o niskiej wierności kopiowania (ang. *error-prone DNA polymerase*), która generuje powstanie zarówno mutacji typu tranzycji jak i transwersji w regionie V genów kodujących immunoglobuliny. Miejsce apirymidynowe może również stać się substratem dla endonukleaz APE1 i APE2 (ang. *apyrimidinic endonuclease*). Katalizują one hydrolizę wiązania fosfodiesterowego w miejscu apirymidynowym, generując powstanie przerwy w jednej z nici DNA. Pierwotne niesparowanie G-U może również zostać rozpoznane przez mechanizmy naprawy niesparowanych zasad (ang. *mismatch repair*, MMR). W tym przypadku nić zawierająca urydynę zostaje częściowo zdegradowana, a polimerazy replikujące uszkodzony DNA takie jak TLS (ang. *translesion synthesis*

polymerase), wypełniają powstałą lukę. Proces ten (tzw. faza 2 SHM) może w konsekwencji prowadzić do licznych tranzycji i transwersji [49].



Rysunek 1.2.2. Rola AID w procesach różnicowania przeciwciał. A) Organizacja *locus* immunoglobulin, w którego skład wchodzi: segmenty VDJ kodujące region zmienny (V) łańcucha ciężkiego immunoglobulin oraz pięć segmentów (C) kodujących region stały łańcucha ciężkiego immunoglobulin. Segmenty C rozdzielone są regionami S (ang. *switch*). B) Proces hiperzmienności somatycznej (SHM). AID generuje mutacje punktowe w obrębie regionu V, indukując w ten sposób proces powstawania przeciwciał specyficznych względem antygeny. Czerwone gwiazdki w regionie V reprezentują mutacje punktowe. C) Proces rekombinacyjnego przełączania klas (CSR). AID inicjuje powstawanie przerw w obu niciach DNA w obrębie regionów S. Inicjowany w ten sposób proces niehomologicznego łączenia końców umożliwia wymianę segmentów C i powstawanie przeciwciał różnych klas. Szczegółowy opis procesów SHM i CSR w tekście. Reprodukacja za zgodą wydawcy © Biotechnology, PAS [51].

Jeśli miejsca apirymidynowe powstaną na obu niciach DNA odpowiednio blisko, możliwe jest powstanie (w fazie 1b SHM) przerwy w obu niciach DNA (ang. *double-stranded DNA breaks*, DSBs). Również mechanizm naprawy MMR, może prowadzić do przekształcenia pęknięć w jednej nici DNA (ang. *single-strand DNA breaks*, SSBs) do DSBs [52]. Inicjowane przez AID powstawanie DSBs (przez deaminację cytydyny do urydyny w obu niciach DNA) w obrębie tzw. regionów S (ang. *switch*) jest niezbędnym warunkiem zajścia drugiego procesu odpowiedzialnego za zmienność przeciwciał – procesu CSR. Regiony S są bogate w powtórzenia tandemowe (AGCT) i są położone w genomie poniżej segmentów kodujących region stały łańcucha ciężkiego immunoglobulin. CSR zachodzi pomiędzy dwoma regionami S przez niehomologiczne łączenie końców (ang. *non-homologous end joining*, NHEJ). NHEJ polega na wewnątrzchromosomowej delecji fragmentu genu pomiędzy dwoma DSBs w regionach S. Proces ten prowadzi w konsekwencji do zamiany fragmentu, kodującego region stały łańcucha ciężkiego (IgH) charakterystyczny dla immunoglobulin klasy M, na inny (wyżej położony) fragment kodujący region stały IgH immunoglobulin klasy G, A lub E. Spekuluje się, że proces ten może zależeć nie tylko od NHEJ, lecz również od rekombinacji homologicznej [53]. Co ważne, segment kodujący region stały IgH

immunoglobuliny klasy D leży pomiędzy tymi samymi regionami S co region stały IgH immunoglobuliny klasy M. Nie dochodzi zatem między nimi do przełączania klas. Pozwala to zrozumieć zjawisko występowania w niepobudzonych limfocytach B równocześnie przeciwciał klasy IgM i IgD (o produkcji jednej lub drugiej klasy decyduje w tym przypadku jedynie proces alternatywnego składania pierwotnego transkryptu) [52, 54]. Schemat obu procesów drugorzędowej zmienności przeciwciał przedstawiony jest na Rysunku 1.2.2.

Opisane powyżej procesy zmienności przeciwciał ukazują, jak fundamentalne znaczenie dla humoralnej odpowiedzi immunologicznej ma stosunkowo prosta, katalizowana przez AID, reakcja deaminacji cytydyny.

1.3. Regulacja ekspresji genu AID

AID ulega ekspresji w indukowanych antygenem limfocytach B w centrum namnażania limfocytów w grudkach limfatycznych. Niski poziom ekspresji *AID* wykryto również we wczesnych stadiach różnicowania limfocytów B w szpiku kostnym, w pierwotnych komórkach płciowych (ang. *primordial germ cells*, PGCs), w embrionalnych komórkach macierzystych (ang. *embryonic stem cells*, ES), w komórkach nabłonkowych prostaty i jajnika oraz w komórkach gruczołu piersiowego [55-57]. W limfocytach B poziom ekspresji *AID* determinuje intensywność procesów SHM i CSR (u myszy o genotypie *AID* +/- procesy te zachodzą na niższym poziomie w porównaniu z myszami *AID*+/+ [58]). Co więcej, fizjologiczna funkcja *AID* jest związana z jej potencjałem onkogennym [59]. Ścisła regulacja ekspresji *AID* jest zatem warunkiem koniecznym zachowania stabilności informacji genetycznej oraz prawidłowości przebiegu procesów immunologicznych.

Ekspresja genu *AID* jest regulowana na wielu poziomach:

1. transkrypcyjnym przez aktywatory i represory transkrypcji;
2. potranskrypcyjnym poprzez regulację stabilności mRNA *AID*;
3. potranslacyjnym poprzez regulację stabilności, lokalizacji i aktywności *AID*;

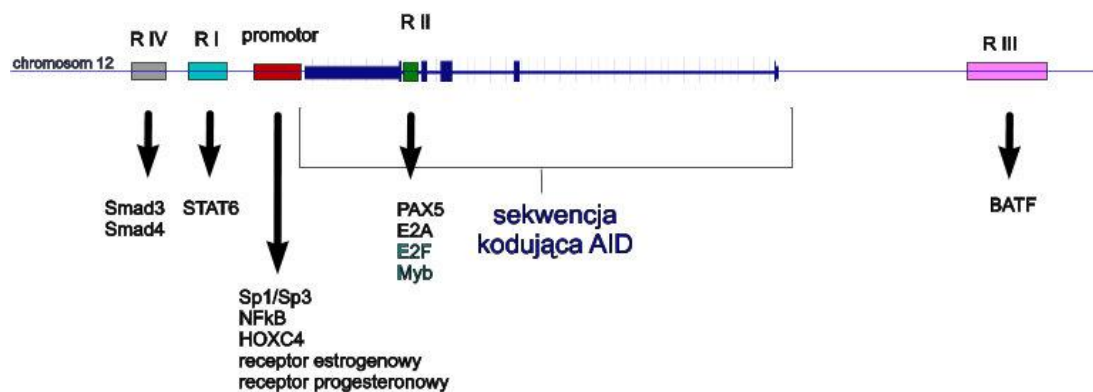
Dodatkowo występuje wiele czynników komórkowych mogących pośrednio lub bezpośrednio kształtować aktywność *AID* i jej specyficzność względem genów *Ig*.

1.3.1. Regulacja transkrypcji

Nasza dotychczasowa wiedza na temat szlaków regulacji ekspresji *AID* pochodzi z badań procesu różnicowania przeciwciał w aktywowanych limfocytach B. W promotorze *AID* oraz w regionach powyżej i poniżej miejsca startu transkrypcji zidentyfikowano miejsca wiązania dla wielu czynników transkrypcyjnych. Sugeruje to, że ekspresja *AID* jest regulowana przez swoistą równowagę pomiędzy aktywatorami i represorami transkrypcji.

W obrębie promotora *AID* znajdują się miejsca wiązania dla stosunkowo powszechnie występujących czynników transkrypcyjnych: Sp1/Sp3, NFκB, HoxC4 oraz miejsca wiązania dla receptorów estrogenowych i progesteronowych [60, 61]. Aktywacja szlaku sygnałowego NFκB zachodzi pod wpływem białek indukujących proces CSR, którymi są: CD40, TLR, BCR, TACI. Białka te mogą aktywować zarówno

kanoniczny jak i niekanoniczny szlak NFκB. Co istotne, należące do niekanonicznego szlaku białko p52 (budujące jeden z heterodimerów NFκB) wiąże się do promotora *AID*, natomiast białko p65, należące do szlaku kanonicznego (również budujące heterodimer NFκB), wiąże się do sekwencji wzmacniających (ang. *enhancer*) zlokalizowanych powyżej promotora. Stymulacja zarówno kanonicznej jak i niekanonicznej ścieżki ma kluczowe znaczenie dla podtrzymania ekspresji *AID* w aktywowanych limfocytach B. Indukcja kanonicznego szlaku gwarantuje szybką, ale zazwyczaj przejściową ekspresję. Szlak niekanoniczny związany jest natomiast ze stałą ekspresją genu [62]. Inny czynnik transkrypcyjny HoxC4 jest aktywowany przez stymulatory różnicowania limfocytów B: CD154, LPS, IL-4, które również biorą udział w indukowaniu ekspresji *AID*. Związanie do promotora różnych czynników transkrypcyjnych daje synergistyczny efekt stymulacji ekspresji *AID*.



Rysunek 1.3.1. Lokalizacja regionów regulatorowych w locus *AID*. Kolorem granatowym zaznaczono sekwencję kodującą *AID* (zlokalizowaną w chromosomie 12) z uwzględnieniem lokalizacji eksonów (pogrubione regiony) i intronów. Promotor genu *AID* oznaczono czerwonym prostokątem. Prostokątami oznaczono także cztery regiony regulatorowe leżące w locus *AID* (od RI do RIV odpowiednio niebieski, zielony, różowy i szary). Poniżej promotora i każdego z regionów regulatorowych wymieniono wiążące się do nich aktywatory i represory transkrypcji (opis w tekście).

Oprócz promotora, locus *AID* zawiera cztery zachowawcze ewolucyjnie regiony regulatorowe (rozciągające się na odcinku od 9 kpz powyżej pierwszego eksonu do 18 kpz poniżej ostatniego eksonu) mające kluczowy wpływ na proces transkrypcji (Rysunek 1.3.1.). Region I zlokalizowany jest powyżej promotora i wiąże czynnik STAT6 (transduktor sygnału oraz aktywator transkrypcji), który jest aktywowany przez IL4. Do regionu II, leżącego w pierwszym intronie, wiązane są stymulatory transkrypcji: PAX5 i E2A. Region III, leżący 17 kpz poniżej promotora, wiąże czynnik

transkrypcyjny BATF. Region IV zlokalizowany jest 9 kbp powyżej promotora i wiąże Smad3 oraz Smad4 (aktywowane przez STAT6 i TGF- β). Wiążące się z regionami regulatorowymi aktywatory transkrypcji współdziałają z czynnikami transkrypcyjnymi NF κ B, HoxC4 i Sp1/Sp3 związanymi do promotora i elementów wzmacniających.

Jak już wspomniano, ekspresja *AID* wydaje się być wypadkową działania skomplikowanych szlaków zarówno aktywacji jak i represji transkrypcji [62, 63]. W regionie regulatorowym II znajdują się cztery elementy *cis* wiążące represory transkrypcji: dwa miejsca wiązania dla Myb, jedno miejsce wiązania dla E2F oraz 350 nt sekwencja bogata w CT. Aktywatory E2A i PAX5, wiążące się w tym samym regionie regulatorowym, znoszą ten efekt inhibicji, ale tylko częściowo. Dopiero wiązanie do sekwencji wzmacniającej czynników: NF κ B, Smad3/4, STAT6 i C/EBP, umożliwia całkowite zniesienie działania represorów intronu pierwszego [63, 64]. Wydaje się zatem, że podstawową przyczyną komórkowo-specyficjnej ekspresji *AID* jest z jednej strony wszechobecność represorów transkrypcji tego genu, z drugiej produkcja unikatowych aktywatorów transkrypcji w stymulowanych cytokinami limfocytach B [57, 65].

Transkrypcja *AID* podlega również regulacji hormonalnej, a estrogen i progesteron działają antagonistycznie w tym procesie. Estrogen stymuluje ekspresję *AID* w limfocytach B (z obserwowanym wzrostem intensywności procesów SHM i CSR), podczas gdy progesteron ma na nią hamujący wpływ. Elementy wiążące progesteron (ang. *progesterone binding element*, PRE) zostały wykryte w obrębie promotora *AID* i zakłada się, że działają one antagonistycznie do zależnego od TNF α szlaku stymulacji transkrypcji *AID* (pozostając jednocześnie bez wpływu na działanie kompleksu aktywacyjnego NF κ B) [66]. Elementy wiążące estrogen (ang. *estrogen response element*, ERE) funkcjonują natomiast niezależnie od szlaku aktywacji *AID* przez czynnik NF κ B i synergistycznie do niego [67]. Sugeruje się, że hormonalna regulacja ekspresji *AID* wspiera humoralną odpowiedź immunologiczną u kobiet, ale jednocześnie implikuje częstsze występowanie u nich zaburzeń autoimmunologicznych i nowotworów w tkankach czułych na działanie estrogenu takich jak gruczoł piersiowy [68].

W regulacji ekspresji *AID*, podobnie jak w regulacji ekspresji wszystkich innych genów, kluczową rolę odgrywają również czynniki epigenetyczne takie jak metylacja DNA czy też metylacja i acetylacja histonów w *locus AID* [69].

1.3.2. Regulacja potranskrypcyjna

Nieulegający translacji region 3'-końcowy (ang. *3' untranslated region*, 3'UTR) mRNA AID jest miejscem wiązania dla dwóch miRNA: miR-155 i miR-181b, o udowodnionym wpływie na stabilność transkryptu AID i co za tym idzie na poziom jego akumulacji w limfocytach B [70, 71]. Zidentyfikowano obecność obu tych miRNA podczas procesu CSR. Mutacja w miejscu wiązania dla miR-155 powoduje 2-3 krotny wzrost poziomu akumulacji mRNA AID, co przekłada się na 1,5-3 krotny wzrost efektywności CSR oraz aż 5 krotny wzrost częstości występowania, związanych z aktywnością AID, translokacji chromosomowych pomiędzy *locus* immunoglobulin i protoonkogenem MYC (tzw. translokacji MYC/IgH). Co ciekawe, nie odnotowano przy tym wzrostu intensywności procesu SHM, co może świadczyć o osobnym mechanizmie regulującym ten proces, w sposób niezależny od ilości transkryptu [70, 72]. Wzrost poziomu miR-181b redukuje intensywność procesu CSR o 50% czemu towarzyszy analogiczny spadek poziomu akumulacji transkryptu i białka AID [71]. Co istotne, w trakcie stymulacji limfocytów B w kierunku CSR, poziom miR-181b spada, podczas gdy poziom miR-155 rośnie. Sugeruje to, że oba miRNA pełnią różne funkcje w regulacji procesów immunologicznych [57]. W pierwotnych liniach limfocytów B transkrypcja AID jest indukowana w pierwszym dniu ich aktywacji, ale w odróżnionych limfocytach B mRNA AID nie jest wykrywana. Prawdopodobnie dzieje się tak zarówno za sprawą inhibicji transkrypcji AID przez czynniki regulatorowe jak i za sprawą zależnego od miRNA potranskrypcyjnego wyciszania genu AID [52, 62].

1.3.3. Regulacja potranslacyjna

Regulacja lokalizacji i stabilności AID. Podstawowe funkcje AID związane są z lokalizacją jądrową. Ponieważ jednak aktywność AID może stanowić zagrożenie dla stabilności genomu białko to zlokalizowane jest głównie w cytoplazmie. AID podlega ciągłemu transportowi pomiędzy jądrem komórkowym i cytoplazmą, przy czym równowaga tego procesu przesunięta jest w stronę eksportu z jądra komórkowego do cytoplazmy. Postuluje się, że niesymetryczność lokalizacji AID jest osiągana przez: aktywne wypompowywanie AID z jądra komórkowego, retencję AID w cytoplazmie oraz wyższy współczynnik degradacji frakcji jądrowej AID w porównaniu z frakcją cytoplazmatyczną [62].

Doniesienia różnych grup badawczych wskazują na istnienie silnego sygnału eksportu jądrowego (ang. *nuclear export signal*, NES) na końcu C białka AID. Nie ma jednak zgodności wśród badaczy co do lokalizacji i natury sygnału lokalizacji jądrowej (ang. *nuclear localization signal*, NLS). Prawdopodobnie jest to słaby sygnał, znajdujący się na końcu N tego białka [62, 73]. Nie jest jasne zatem, w jaki sposób AID jest transportowana do jądra komórkowego. Jest to na tyle małe białko, że mogłoby dyfundować przez pory w otoczce jądrowej. Sugeruje się również możliwość oddziaływania AID z szeregiem izoform importyny A. Nieco więcej wiadomo na temat transportu AID z jądra komórkowego do cytoplazmy. Transport ten odbywa się na zasadzie aktywnego eksportu zależnego od receptora Xpo1 (eksportyna 1). Komórki hodowane w obecności leptomycyny B (specyficzny inhibitor Xpo1) wykazują wyższy poziom akumulacji AID w jądrze komórkowym. Co ciekawe, mutacje w obrębie NES, prowadzą do zahamowania procesu CSR lecz nie wpływają na proces SHM [74]. Oznacza to, że maszyneria eksportu jądrowego jest ściśle związana z aktywnością AID w procesie CSR. Molekularne podstawy tej zależności pozostają jednak nieznanne.

Znacząca dla lokalizacji AID jest niestabilność jej frakcji jądrowej w porównaniu z frakcją cytoplazmatyczną. Okres półtrwania AID w jądrze komórkowym wynosi zaledwie 2,5 h, podczas gdy dla frakcji cytoplazmatycznej wynosi on 18 h. Przyjmuje się, że poliubikwitynacja AID i jej degradacja w proteasomie zachodzi dużo intensywniej w jądrze komórkowym [75]. Ponadto wykazano, że białko szoku cieplnego Hsp90 chroni frakcję cytoplazmatyczną przed poliubikwitynacją i degradacją, zapewniając tym samym cytoplazmatyczny rezerwuar AID [76]. W cytoplazmie AID oddziałuje z szeregiem białek opiekuńczych tworząc kompleks o względnie dużej masie molekularnej, zwany cytoplazmatycznym kompleksem AID. Oprócz wspomnianego Hsp90 w skład tego kompleksu wchodzi również białka towarzyszące białkom opiekuńczym: Hsp40, Hsp70, DnaJa1 oraz translacyjny czynnik elongacyjny eEF1A. Cytoplazmatyczny kompleks AID warunkuje zatem cytozolową retencję i stabilność AID.

Choć mechanizmy molekularne regulujące równowagę pomiędzy jądrową i cytoplazmatyczną lokalizacją AID nie są w pełni poznane, to wiadomo że zabezpieczają one genom limfocytów B przed mutagennym działaniem AID oraz zapewniają cytoplazmatyczny rezerwuar tego białka niezbędny do szybkiej odpowiedzi immunologicznej.

Modyfikacje potranslacyjne a regulacja aktywności AID. Jediną stwierdzoną modyfikacją potranslacyjną jakiej podlega AID jest fosforylacja. W sekwencji aminokwasowej tego białka znaleziono wiele potencjalnych miejsc fosforylacji: S3, T27, S38, T140, Y184. Niemniej jednak w oczyszczonym białku AID izolowanym z aktywowanych limfocytów B eksperymentalnie (metodą spektrometrii mas) potwierdzono obecność modyfikacji tylko na S38, T140 i Y184. Nie odnotowuje się wpływu tych fosforylacji na aktywność AID *in vitro* lub podczas badania aktywności AID w komórkach *E. coli*. Mają one natomiast zasadniczy wpływ na przebieg procesów SHM i CSR w warunkach *in vivo*.

W aktywowanych limfocytach B jedynie niewielka frakcja AID jest fosforylowana w pozycji S38, ale to właśnie ta frakcja preferencyjnie wiąże się z chromatyną. Modyfikacja S38 przez kinazę białkową A (ang. *protein kinase A*, PKA), warunkuje oddziaływanie AID z białkiem RPA (ang. *replication protein A*). RPA wiąże ssDNA i bierze udział w kierowaniu AID do miejsc docelowych w genomie. Mutacja S38A redukuje intensywność procesów SHM i CSR aż o 80-90% w porównaniu z AID typu dzikiego. Ponadto, w systemie transkrypcji *in vitro* obecność białka RPA zwiększa aktywność AID, a mutacje reszty S38 w sekwencji aminokwasowej AID zaburzają oddziaływanie pomiędzy tymi dwoma białkami. W warunkach *in vivo*, regiony aktywne transkrypcyjnie mogą być efektywnie deaminowane tylko przez fosforylowaną w tej pozycji formę AID. Co więcej, PKA może się wiązać z regionami S w sposób niezależny od AID [77]. Sugeruje to, że aktywacja AID przez PKA (i powstanie formy AID zdolnej do oddziaływania z RPA) następuje dopiero w pobliżu sekwencji docelowej. Ukazuje to, jak ważną rolę pełni fosforylacja S38 i oddziaływanie z białkiem RPA podczas regulacji aktywności i wiązania AID do ssDNA [78-80].

Fosforylacja reszty T140 wydaje się mieć dużo mniejszy wpływ na aktywność AID. Mutacja w tym miejscu redukuje intensywność procesu SHM (o 60% w porównaniu z AID typu dzikiego) i nie ma znaczącego wpływu na proces CSR [77, 81]. Mechanizm, na drodze którego fosforylacja ta wpływa na przebieg procesu SHM nie jest znany. Prawdopodobnie fosforylacja tej reszty warunkuje oddziaływanie AID z jeszcze nieznanym kofaktorem, specyficznym dla procesu SHM. Znaczenie funkcjonalne fosforylacji Y184 nie jest znane [62].

1.4. Specyficzność AID względem genów immunoglobulin

AID sporadycznie deaminuje inne miejsca w genomie niż *locus Ig*. Taka niespecyficzna aktywność AID może prowadzić do destabilizacji materiału genetycznego i rozwoju nowotworów. Stąd, jednym z najbardziej interesujących pytań, przed którym stoi współczesna immunologia jest to w jaki sposób AID rozpoznaje *locus Ig*? Na podstawie dotychczasowych obserwacji można zakładać, że selektywność działania AID jest osiągana zarówno przez czynniki specyficzne jak i niespecyficzne. Determinują ją: sekwencja wiązanych regionów genomu, ich aktywność transkrypcyjna, czynniki *cis* i *trans* oraz epigenetyczne modyfikacje chromatyny. Selektywność działania AID jest również osiągana w wyniku kooperacji AID z szeregiem kofaktorów białkowych.

AID działa w czasie transkrypcji, preferencyjnie w obrębie motywów WRCY. Deaminacja cytydyny do urydyny w obrębie regionów V lub S genów *Ig* inicjuje, odpowiednio, proces SHM lub CSR. Analiza mutacji generowanych przez AID zarówno w warunkach *in vitro* jak i *in vivo* wykazała, że AID preferencyjnie działa na cytydyny w obrębie specyficznego kontekstu sekwencyjnego WRCY (W = A/T; R = A/G; Y = C/T). Co ważne, mutacje w regionach V immunoglobulin pojawiają się 100-200 pz poniżej miejsca startu transkrypcji stopniowo zanikając na odcinku 1,5-2 kpz (jeszcze przed osiągnięciem intronowej sekwencji wzmacniającej i regionu C). W procesie CSR AID inicjuje powstawanie mutacji w regionach S około 150 pz poniżej miejsca inicjacji transkrypcji, na odcinku 2-12 kpz. Zarówno w procesie SHM jak i CSR deaminowane są obie nici DNA [82, 83]. W przeciwieństwie do regionów V, które nie różnią się pod względem składu nukleotydowego od innych miejsc w genomie, regiony S są bogate w pary GC, a ich sekwencja składa się z powtórzeń AGCT (będących preferowanymi przez AID motywami) [84].

Ponieważ substratem w reakcji deaminacji katalizowanej przez AID jest ssDNA, aktywność AID jest ściśle związana z aktywnością transkrypcyjną rozpoznawanych regionów DNA. W rezultacie procesom SHM i CSR ulegają jedynie aktywne transkrypcyjnie (po rearanzacji VDJ) geny immunoglobulin, a aktywność transkrypcyjna tych genów silnie koreluje z aktywnością AID [85, 86]. AID deaminuje ssDNA w powstającym podczas transkrypcji tzw. bąblu transkrypcyjnym (długości kilkunastu nukleotydów). Co ciekawe, w badaniach *in vitro* wykazano, że aktywność AID w bąblu transkrypcyjnym jest trzykrotnie wyższa niż wobec ssDNA [32]. Obecnie

wiadomo, że AID deaminuje cytydynę na obu niciach regionów S i regionu V z porównywalną wydajnością, chociaż obserwowana jest nieznaczna preferencja AID względem nietranskrybowanej nici [49, 83]. Istnieje hipoteza tłumacząca dostęp AID do obu nici DNA poprzez tworzenie się ujemnych skrętów superhelikalnych. Według tej hipotezy AID może rozpoznawać dwuniciowe odcinki DNA pod warunkiem ich superhelikalnego zwinięcia. W takich superhelikalnych skrętach AID wykazuje jednakowe powinowactwo do obu nici DNA. Skręty te tworzą się na odcinkach DNA powyżej kompleksu transkrypcyjnego więc ich powstawanie jest ściśle związane z transkrypcją. Hipoteza ta stoi jednak w sprzeczności z powszechnie panującym przekonaniem, iż enzym ten działa wyłącznie na ssDNA. [87]. Innym proponowanym mechanizmem tłumaczącym dostęp AID do obu nici DNA jest formowanie się w obrębie bogatego w pary GC regionu S charakterystycznych pętli R (ang. *R-loops*). Ich obecność powoduje ekspozycję dłuższego niż w typowym bąblu transkrypcyjnym jednoniciowego odcinka DNA [88]. Niektórzy badacze sugerują ponadto, iż AID uzyskuje dostęp do transkrybowanej nici poprzez oddziaływanie z egzosomem [89]. Egzosom jest kompleksem białkowym zaangażowanym w kontrolę jakości nowo powstającego RNA i ma zdolność jego degradacji. Kompleks egzosomu może łączyć się z kompleksem polimerazy RNA II i usuwać tworzący się transkrypt. Proces ten prowadzi do ekspozycji nici transkrybowanej i udostępnienia jej dla AID [90]. W systemie transkrypcji *in vitro* obecność białek tworzących egzosom umożliwia zajście deaminacji na obu niciach DNA w bąblu transkrypcyjnym [89].

Jak wspomniano w poprzednim podrozdziale, do zajścia procesów SHM i CSR konieczne jest oddziaływanie AID z białkiem RPA wiążącym jednoniciowy DNA. AID może również oddziaływać z kompleksem polimerazy RNA II do którego jest rekrutowana za pośrednictwem białka Spt5. Jak pokazują wyniki badań *in vitro*, Spt5 bezpośrednio oddziałuje z AID. Kolokalizację AID, Spt5 i kompleksu polimerazy RNA II potwierdzono metodą ChIP-seq (ang. *chromatin immunoprecipitation followed by massive parallel sequencing*), a wyciszenie genu *Spt5* skutkuje inhibicją wiązania AID do regionów S [91]. Przypuszcza się, że Spt5 stabilizuje jednoniciowe odcinki DNA, jest czynnikiem indukującym pauzy kompleksu transkrypcyjnego oraz bierze udział w oddziaływaniu z kompleksem egzosomu. W ten sposób białko to stymuluje wiązanie AID do transkrybowanej nici DNA w bąblu transkrypcyjnym.

Czynniki *cis* i *trans* odpowiedzialne za wiązanie AID do genów *Ig*. Zaangażowanie czynników *cis* (będących elementem cząsteczki docelowej) w wiązanie AID jest stosunkowo trudne do stwierdzenia, ze względu na ich bezpośredni wpływ na wydajność transkrypcji (każda zmiana sekwencji mająca wpływ na transkrypcję przekłada się na aktywność AID). Pomimo tego, udowodniono wpływ na intensywność przebiegu procesów SHM i CSR takich czynników *cis* jak wzmacniacz intronowy (ang. *intrinsic enhancer*, iE) oraz wzmacniacz intronowy 3' (3'E) [83, 92]. Na podstawie danych literaturowych można również sądzić, że motyw CAGGTG, będący miejscem wiązania dla białek rozpoznających tzw. kasetę E (takich jak E47) ma bezpośredni wpływ na wiązanie AID pozostając jednocześnie bez wpływu na wydajność transkrypcji. Mechanizm zależnego od kasety E wiązania AID nie jest do końca znany. Na jego istotną rolę wskazuje jednak fakt, iż kasety te są charakterystyczne również dla innych genów niż geny *Ig*, które są mutowane przez AID w sposób niespecyficzny [93, 94].

AID nie wydaje się być w stanie samodzielnie rozpoznawać sekwencji regulatorowych. Sekwencje regulatorowe, specyficzne dla genów *Ig*, są rozpoznawane przez szereg czynników *trans*, za których pośrednictwem AID jest rekrutowana do regionów V i S genów *Ig*. Udowodniono oddziaływanie AID ze wspomnianymi już białkami: PKA, RPA, UDG, z białkami kompleksu naprawy niesparowań MSH2/MSH6, kompleksu polimerazy RNA II (poprzez białko Spt5) i kompleksu egzozomu. Ponadto AID oddziałuje z represorem transkrypcji KAP1/Trim28, kinazą białkową zależną od DNA DNA-PK (zaangażowaną w proces NHEJ), białkiem 14-3-3 rozpoznającym sekwencje powtórzone AGCT w regionach S, białkiem PTBP2 hamującym składanie genów podczas dojrzewania mRNA, oraz innymi białkami zaangażowanymi w dojrzewanie mRNA takimi jak: SRSF1, CTNBL1, GANP. Oddziaływanie AID z białkami SRSF1 i GANP jest specyficzne dla procesu SHM. Natomiast białka 14-3-3, UDG, MSH2/MSH6, PTBP2 oddziałują z AID specyficznie w regionach S [57, 83, 84, 95, 96].

Pomimo że procesy SHM i CSR przebiegają w tych samych komórkach, na tym samym etapie odpowiedzi immunologicznej, dotyczą tego samego *locus* i oba są indukowane przez AID, to mogą one zachodzić zupełnie niezależnie (możliwa jest zajście jednego procesu przy braku drugiego) i być regulowane przez oddzielne szlaki sygnałowe. Za takim poglądem przemawia kilka faktów. Po pierwsze w pierwotnych liniach limfocytów B dużo łatwiej jest aktywować proces CSR, podczas gdy proces

SHM nie występuje wcale lub też zachodzi na bardzo niskim poziomie. Po drugie, powszechna u pacjentów z zespołem HIGM2 delecja odcinka kodującego 9 reszt aminokwasowych z końca C białka AID, wiąże się z całkowitą utratą zdolności przełączania klas przy jednoczesnym zachowaniu zdolności do generowania hiperzmienności somatycznej. Na podstawie wyników badań mutantów AID sugeruje się zatem, że kofaktory białkowe specyficzne dla procesu CSR oddziałują z końcem C białka AID [97]. Również na podstawie wyników badań mutantów przypuszcza się, że koniec N białka AID jest zaangażowany w wiązanie kofaktorów specyficznych tylko dla procesu SHM (mutacje w tym regionie redukują intensywność procesu SHM, ale nie zaburzają przebiegu procesu CSR) [98]. Krótką charakterystykę wybranych mutantów AID z uwzględnieniem wpływu poszczególnych mutacji na procesy SHM i CSR zawiera Tabela 1.4.1.

Wpływ epigenetycznych modyfikacji chromatyny na wiązanie AID do genów *Ig*.

Epigenetyczne modyfikacje chromatyny mają kluczowe znaczenie zarówno w kierowaniu AID do miejsc docelowych, jak i kierowaniu do tych miejsc kofaktorów białkowych oddziałujących z AID. Potranslacyjne modyfikacje histonów takie jak: trimetylacja lizyny 4 histonu H3 (H3K4me3) oraz acetylacja lizyny 9 połączona z fosforylacją seryny 10 w histonie H3 (H3K9acS10ph), są charakterystyczne dla regionów S. Modyfikacje te zapewniają nie tylko fizyczną dostępność chromatyny dla AID i innych białek, ale są również specyficznym znacznikiem i kolejnym poziomem zabezpieczeń gwarantujących specyficzność działania zarówno AID jak i całej maszynarii CSR (dla przykładu, przy braku H3K4me3 nie powstają dwuniciowe przerwy w DNA podczas procesu CSR). Dla prawidłowości procesu SHM kluczowa jest natomiast hipometylacja DNA regionu V *Ig* i specyficzna dla regionów ulegających SHM fosforylacja seryny 14 histonu H2B (H2BS14ph) [62, 69].

Współczesna nauka nie znalazła jak dotąd pełnej i jednoznacznej odpowiedzi na pytanie w jaki sposób osiągana jest specyficzność działania AID względem *locus Ig*. Przedstawione powyżej dane uwidaczniają złożoność tego zjawiska. Z pewnością pełniejsza charakterystyka mechanizmów molekularnych leżących u podłoża aktywności AID przyczyni się do lepszego zrozumienia tej kwestii.

Tabela 1.4.1. Wybrane mutanty AID i ich wpływ na intensywność procesów CSR i SHM.

Lp.	Mutant AID	Lokalizacja mutacji w strukturze białka	Intensywność procesu CSR lub SHM				źródło
			Dane kliniczne		Dane z systemu hodowli komórkowych (% względem typu dzikiego)		
			CSR	SHM	CSR	SHM	
1.	R24W	Miejsce aktywne	-	-	1%*	1%*	[99]
2.	H56Y	Domena deaminazowa	-	-	2%*	1%*	
3.	W68X	Domena deaminazowa	-	-	2%*	1%*	
4.	W80R	Domena deaminazowa	-	-	4%*	1%*	
5.	M139V	Domena charakterystyczna dla białek AID/APOBEC	-	-	5%*	1%*	
6.	Δ189-198	NES	-	+	7%*	75%*	
7.	R112H	Łącznik	-	-	2%	1%	[100]
8.	C147X	Domena charakterystyczna dla białek AID/APOBEC	-	-	2%	0%	
9.	181/182 insercja 34aa	Domena charakterystyczna dla białek AID/APOBEC	-	+	3%	71%	
10.	G23S	Miejsce aktywne	N/A	N/A	97% [°]	11% [°]	[101]
11.	W20K	Miejsce aktywne	N/A	N/A	102%	17%	[98]
12.	D45A	Miejsce aktywne	N/A	N/A	30%	0%	[35]
13.	R50A	Miejsce aktywne	N/A	N/A	130%	25%	
14.	N51A	Miejsce aktywne	N/A	N/A	58%	2%	
15.	K52A	Miejsce aktywne	N/A	N/A	130%	15%	

N/A – niebadane; X – kodon STOP, zielone tło – mutanty zidentyfikowane u pacjentów z zespołem HIGM2; szare tło – mutanty wygenerowane na drodze inżynierii genetycznej; (-)/(+) - brak/obecność procesu SHM lub CSR; (*) dane pochodzące z [100]; (°) dane pochodzące z [98].

1.5. AID a aktywna demetylacja genomu

Białko AID zostało zidentyfikowane jako enzym selektywnie produkowany w aktywowanych antygenem limfocytach B. Najnowsze dane literaturowe wskazują jednak, że AID w warunkach fizjologicznych może również pełnić funkcje poza układem odpornościowym. Sugeruje się, że AID (a także inne białka z rodziny AID/APOBEC) dzięki zdolności deaminacji modyfikowanej cytydyny, bierze udział w aktywnej demetylacji genomu. W ten sposób AID może mieć bezpośredni wpływ na proces reprogramowania epigenetycznego, regulację ekspresji genów i zjawisko pluripotencji komórek.

1.5.1. Metylacja i demetylacja genomu

Metylacja genomu jest jednym z najistotniejszych czynników epigenetycznych determinujących ekspresję genów. Najczęstszą modyfikacją jest metylacja węgla w pozycji 5 cytozyny, co prowadzi do powstania 5-metylocytozyny. Ze względu na powszechność występowania tej modyfikacji oraz jej znaczenie, 5-metylocytydina (5mC) nazywana jest również piątym nukleotydem. Szacuje się, że 70-80% cytydyn w obrębie motywów CpG jest zmetylowane w zróżnicowanych komórkach ssaków. Hipometylowane wyspy CpG obejmują regiony promotorowe większości genów konstytutywnych oraz ok. 40% genów tkankowo-specyficznych [102, 103]. Modyfikacja ta jest zatem podstawowym sposobem represji transkrypcji genów. Jest ona również krytyczna dla stabilności genomu, różnicowania komórek, odpowiada za inaktywację chromosomu X, mechanizm piętnowania rodzicielskiego oraz wyciszenie retrotranspozonów [104, 105]. Obserwuje się również zmiany profilu metylacji w procesach nowotworzenia [106]. U ssaków zidentyfikowano szereg enzymów odpowiedzialnych za metylację cytydyny *de novo* i utrzymywanie metylacji podczas replikacji DNA. Za wprowadzenie metylacji podczas replikacji oraz za jej utrzymanie odpowiadają metylotransferazy DNA (ang. *DNA methyltransferase*, DNMT) (głównie białko DNMT1), natomiast za metylację *de novo* odpowiadają białka DNMT3A oraz DNMT3B [107]. Znacznie mniej wiadomo na temat mechanizmów demetylacji. Proces ten może zachodzić zarówno w sposób bierny jak i aktywny. Bierna demetylacja genomu zachodzi podczas replikacji DNA, kiedy to przy braku aktywności metylotransferaz DNA, niemetylowane reszty cytydyny są włączane do nowo powstałej nici DNA. Mechanizm ten nie tłumaczy jednak zmian w profilu metylacji

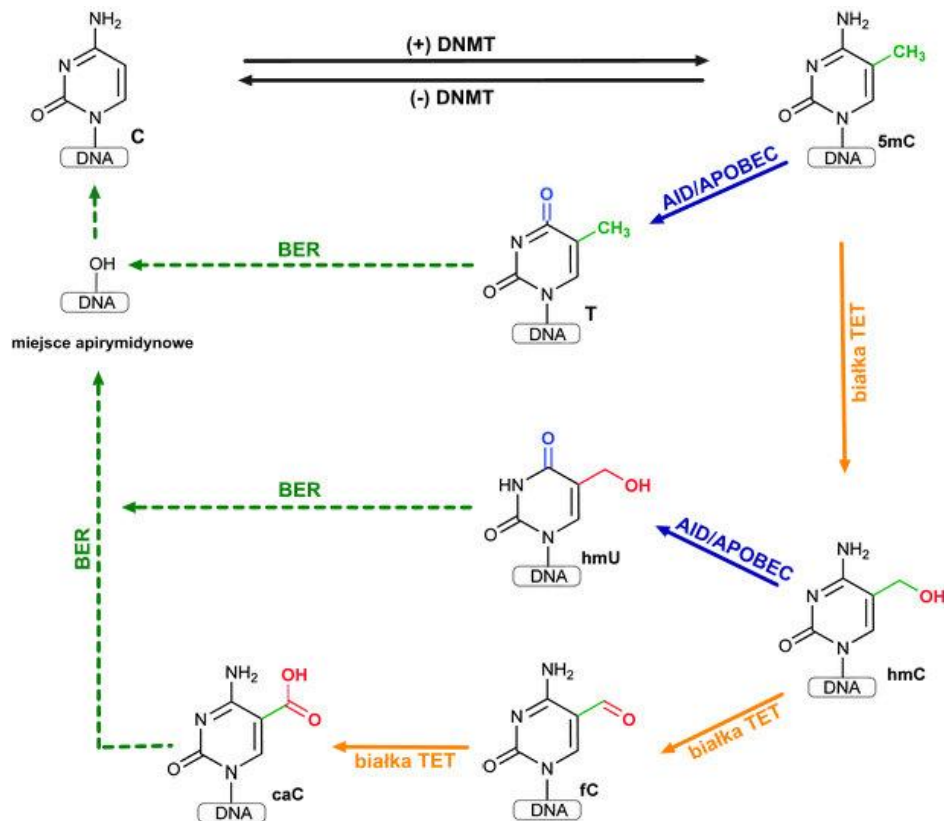
w komórkach, które nie ulegają podziałom. W komórkach takich demetylacja zachodzi w sposób aktywny. Jest to proces enzymatyczny, w którym grupy metylowe są usuwane z DNA w sposób niezależny od replikacji. Aktywna, globalna demetylacja genomu zachodzi na dwóch etapach rozwoju ssaków: w genomie ojcowskim we wstępnej fazie procesu zapłodnienia (przed zlaniem się jąder i pierwszym podziałem zygoty) oraz w trakcie formowania pierwotnych komórek płciowych (ang. *primordial germ cells*, PGCs) między 11,25 a 13,5 dniem rozwoju embrionalnego [108, 109]. Aktywna globalna demetylacja powoduje przeprogramowanie genomu i powrót komórek do stanu pluripotencji. Jest zatem niezbędna w trakcie przebiegu procesów różnicowania i rozwoju [108, 110-112]. Część badaczy sugeruje, że proces aktywnej demetylacji może występować również na dalszych etapach rozwoju, powodując zmiany ekspresji jedynie poszczególnych genów w komórkach zróżnicowanych. Taka specyficzna wobec *locus* aktywna demetylacja jest niezbędna do reaktywacji genów pluripotencji podczas reprogramowania komórki [113]. Występowanie tego procesu w komórkach zróżnicowanych wciąż pozostaje kontrowersyjne, głównie ze względu na problemy doświadczalne, których następcza oddzielenie efektów związanych z procesami aktywnej i biernej demetylacji oraz spontanicznej deaminacji 5mC. Niemniej jednak, sugeruje się udział aktywnej demetylacji w regulacji ekspresji genów m.in. podczas różnicowania monocytów w makrofagi i komórki dendrytyczne [114] oraz podczas regulacji ekspresji *IL-2* w limfocytach T [115].

1.5.2. Proponowane mechanizmy aktywnej demetylacji

Obecnie niewiele wiadomo na temat mechanizmów aktywnej demetylacji genomu u ssaków. Postuluje się, że jest to wieloetapowy proces angażujący mechanizmy naprawy DNA oraz modyfikacje 5mC do innych związków. U roślin kwiatowych takich jak *Arabidopsis thaliana* demetylację DNA przeprowadzają glikozylazy 5-metylocytydiny takie jak DEMETER lub ROS, które specyficznie usuwają 5-metylocytydynę (bez uprzedniej modyfikacji) pozostawiając miejsce apirymidynowe (ang. *apyrimidinic site*, AP). Miejsce AP jest substratem dla mechanizmów naprawy mutacji takich jak mechanizm naprawy przez wycięcie zasad (ang. *base excision repair*, BER) [116]. U ssaków taka klasa 5-metyloglikozylaz wydaje się nie istnieć [117], dlatego poszukiwane są alternatywne drogi demetylacji. Wstępnie proponowano białko MBD4 jako odpowiednik tych enzymów u ssaków, jednakże

obecnie uważa się, że jest ono dużo bardziej aktywne względem tymidyny niż 5mC (nawet trzydziestokrotnie) [118].

Mechanizmy leżące u podstawy aktywnej demetylacji u zwierząt są źródłem wielu kontrowersji. Proponowanych jest kilka potencjalnych ścieżek. W najprostszej sytuacji grupa metylowa przy węglu 5 mogłaby być po prostu usunięta. Rodzina enzymów Alkb ma zdolność demetylowania 1-metyloadenozyny i 3-metylocytydyny poprzez utleniony produkt pośredni, uwalniając węgiel z grupy metylowej jako formaldehyd i regenerując oryginalną zasadę azotową [119]. U kręgowców jednak, jak do tej pory, nie został zidentyfikowany enzym z tej rodziny wykorzystujący 5mC jako substrat. W rezultacie, proponowane obecnie mechanizmy aktywnej demetylacji u ssaków zakładają wcześniejszą modyfikację 5mC i następujące po niej usunięcie zmodyfikowanej zasady azotowej przez szlaki naprawy mutacji takie jak BER. Postulowane są trzy możliwe szlaki modyfikacji 5mC (Rysunek 1.5.1). Pierwszy polega na bezpośredniej deaminacji 5mC do T przez deaminazy takie jak rodzina białek AID/APOBEC (tzw. demetylacja zależna od deaminacji). Deaminacja 5mC generuje w tym przypadku niesparowanie T-G, rozpoznawane i naprawiane przez system BER (głównie glikozylazę TDG i MBD4). Drugi proponowany szlak demetylacji DNA przebiega poprzez oksydację 5mC do 5hmC (5-hydroksymetylocytydyny) katalizowaną przez białko TET1, a następnie deaminację 5hmC do 5hmU (5-hydroksymetylourydyny) rozpoznawanej przez szlaki naprawy mutacji (ponownie, głównie glikozylazę TDG i MBD4) [120, 121]. Wyniki badań Guo i współpracowników świadczą, że transfekcja embrionalnych komórek nerki wektorami zawierającymi sekwencje kodujące białka TET1 i AID obniża poziom genomowej 5hmC. Stwierdzono również obniżenie genomowego poziomu 5hmC podczas transfekcji wektorami zawierającymi sekwencje kodujące białka APOBEC1, APOBEC2, APOBEC3A, APOBEC3C, APOBEC3E, ale nie podczas transfekcji wektorami zawierającymi sekwencje kodujące białka APOBEC3G i APOBEC3B [122]. Niemniej jednak wyniki badań *in vitro* przeczą temu, by AID/APOBEC wykazywały zdolność deaminacji 5hmC [123, 124]. W ostatnim z postulowanych mechanizmów aktywnej demetylacji, 5hmC powstała w trakcie katalizowanej przez białko TET oksydacji 5mC, ulega dalszej (kilkuetapowej) oksydacji do 5fC (5-formylocytydyny) i 5caC (5-karboksylocytydyny) (tzw. demetylacja niezależna od deaminacji). Te analogi cytydyny o wyższym stanie oksydacyjnym zostały wykryte w genomie embrionalnych komórek macierzystych (ang. *embryonal stem cells*, ES) i są dobrymi substratami dla enzymu TDG [125, 126].



Rysunek 1.5.1. Proponowane mechanizmy demetylacji 5mC u ssaków. Proces demetylacji może przebiegać w sposób bierny jak i aktywny. O biernej demetylacji DNA mówimy, gdy podczas replikacji DNA brak jest aktywności metylotransferaz DNA (DNMT, czarne strzałki na schemacie). Aktywna demetylacja DNA jest procesem wieloetapowym, zakładającym modyfikację 5mC do innych związków, które mogą być rozpoznane przez mechanizmy naprawy mutacji, głównie BER (zielone strzałki). Obecnie postulowane są trzy alternatywne szlaki modyfikacji 5mC: (1) deaminacja 5mC do tymidyny przez białko(a) z rodziny AID/APOBEC; (2) oksydacja 5mC kolejno do 5hmC, 5fC i 5caC przez białka z rodziny TET; (3) oksydacja 5mC do 5hmC przez białka z rodziny TET, a następnie deaminacja 5hmC do 5hmU przez białka z rodziny AID/APOBEC. Na podstawie [127, 128].

Wśród badaczy brak jest zgody co do tego, czy za aktywną demetylację odpowiada jeden, czy też kilka potencjalnych mechanizmów, oraz czy mechanizm ten jest wspólny dla różnych grup zwierząt. Nie ma jednak wątpliwości, że zarówno u roślin jak i u zwierząt naprawa mutacji odgrywa kluczową rolę w procesie aktywnej demetylacji.

1.5.3. Udział AID w procesach aktywnej demetylacji

Od czasu zidentyfikowania ekspresji genu AID w komórkach PGC i na wczesnych etapach rozwoju embrionalnego niższych kręgowców sugeruje się, że białko to może mieć związek z procesami demetylacji. Ze względu na pełnioną funkcję w procesach różnicowania przeciwciał, przyjęto uważać, że AID ulega ekspresji jedynie w aktywowanych limfocytach B. Dlatego też ogromnym zaskoczeniem było wykrycie w 2004 roku ekspresji tego genu poza układem odpornościowym – w komórkach pluripotencjalnych takich jak: oocyty, komórki PGC i ES [55]. W kolejnych badaniach pokazano również ekspresję AID w komórkach epitelialnych prostaty [129] i w gruczole piersiowym [67]. Obserwowana ekspresja, choć na niższym poziomie niż w aktywowanych limfocytach B, była na tyle znacząca, że pozwalała zasugerować potencjalną funkcję AID poza układem immunologicznym. Stwierdzona w dalszych badaniach obecność AID we wczesnych stadiach embriogenezy *Danio rerio* [130] i *Xenopus leavis* [131] dodatkowo przemawia za udziałem AID w mechanizmach rozwojowych.

Po raz pierwszy zdolność AID do katalizowania reakcji deaminacji 5mC do T pokazano w 2004 roku. Częściowo oczyszczone ludzkie białko AID i szczurze APOBEC1 wykazywały zdolność deaminacji 5mC w DNA zarówno *in vitro* jak i w komórkach *E. coli*. Bazując na tych obserwacjach oraz na stwierdzonej ekspresji AID w komórkach pluripotencjalnych, badacze wysunęli hipotezę, iż AID przez deaminację 5mC do T może brać udział w procesie aktywnej demetylacji genomu i w ten sposób uczestniczyć w procesach reprogramowania [55]. Zgodnie z zaproponowanym mechanizmem, powstałe w wyniku działania AID niesparowanie G-T, powinno zostać rozpoznane przez białka odpowiedzialne za naprawę mutacji takie jak białka szlaku BER. Ich działanie skutkowałoby wstawieniem niemetylowanej cytydyny w miejscu pierwotnego niesparowania. W ten sposób maszyna naprawy DNA usuwałaby znacznik epigenetyczny w procesie inicjowanym przez AID.

Dalsze dowody na udział AID w procesach demetylacji przyniosły badania na embrionach *Danio rerio* [130]. Badacze zauważyli, że wprowadzenie metylowanej cząsteczki DNA do pojedynczej komórki embrionu indukuje ekspresję AID oraz APOBEC2. Co więcej stwierdzono, że podwyższona ekspresja AID prowadzi do efektywnej demetylacji wprowadzonej cząsteczki DNA (z obserwowaną konwersją 5mC do T). Ponadto dodatkowa transfekcja genem glikozylazy *MBD4* zwiększa ten efekt. W zaproponowanym przez badaczy mechanizmie aktywnej demetylacji

u niższych kręgowców, AID deaminuje 5mC do T, która to następnie jest rozpoznawana i usuwana przez białko MBD4, rekrutujące również inne białka maszynerii BER. W mechanizmie tym białko Gadd45 służy jako łącznik pomiędzy AID i MBD4. Wyciszenie ekspresji *AID*, *MBD4* lub *Gadd45* wiązało się ze wzrostem poziomu metylacji genomu *Danio rerio*. Ponadto wyciszenie ekspresji *AID* lub *MBD4* prowadziło do hipermetylacji promotora genu *NEUROD2* (uważanego za jeden z genów pluripotencji), co może sugerować bezpośredni udział tych białek w procesach rozwojowych i w reprogramowaniu genomu [130].

Brak jest dowodów na zaangażowanie białka Gadd45 w demetylację u ssaków. Co więcej mutant *Gadd45* *-/-* nie wykazuje znaczących różnic w poziomie metylacji [132]. Istnieją jednak dowody na udział AID w tym procesie. W mysich komórkach PGC *AID* *-/-* występuje znacząco wyższy poziom globalnej metylacji genomu [133]. Niemniej jednak komórki te, pomimo braku ekspresji *AID*, utrzymują pewien stopień demetylacji. Nie stwierdza się również znaczących zaburzeń rozwojowych u myszy *AID* *-/-* [134]. Także osoby z zespołem HIGM2 nie wykazują upośledzeń poza układem odpornościowym [48]. Nie jest jednak jednoznaczne czy utrata funkcji jednego z białek zaangażowanych w proces demetylacji musiałaby skutkować drastycznymi zmianami fenotypu. U *A. thaliana* utrata funkcji wszystkich trzech znanych glikozylaz specyficznych względem 5mC nie ma znacznego odbicia w fenotypie (rośliny są w pełni żywotne, a hipermetylowane są tylko niektóre *loci*) [135]. Można zatem przypuszczać, że oprócz AID również inne czynniki są zaangażowane w proces aktywnej demetylacji u ssaków, a AID pełni bardziej subtelną funkcję w demetylacji tylko niektórych regionów genomu. W komórkach powstałych przez fuzję mysich ES z ludzkimi fibroblastami (tzw. system heterokarionowy) wykazano wpływ ekspresji *AID* na poziom metylacji regionów promotorowych genów pluripotencji – *NANOG* i *OCT4*. Wyciszenie ekspresji *AID* w tych komórkach powodowało wtórną metylację genów *NANOG* i *OCT4* i w efekcie ich unieczynnienie [113]. Wyniki te przemawiają za udziałem AID w utrzymywaniu i/lub generowaniu stanu pluripotencji oraz w procesach odróżnicowywania.

Pojawiają się również głosy sceptyczne, podważające udział AID w procesach epigenetycznych u ssaków. W 2010 roku Hejkova i współpracownicy pokazali, iż aktywna demetylacja w pierwotnych komórkach płciowych jest silnie związana z aktywacją szlaków BER, ale nie ma związku z ekspresją *AID*. Stwierdzono także całkowity brak ekspresji *AID* w mysich komórkach PGC w 11.5 dniu embriogenezy

oraz niewielki poziom ekspresji *AID* w 12.5 dniu (epigenetyczne procesy reprogramowania zaczynają się w 11.5 dniu) [136]. Biorąc jednak pod uwagę fakt, iż dostępnymi systemami do badań wczesnych mechanizmów embriogenezy oraz procesów aktywnej demetylacji są tak skomplikowane układy jak embriony niższych kręgowców, mysie komórki PGC, ludzkie ES czy też system heterokarionowy, nie dziwi pewna rozbieżność otrzymywanych wyników. Dlatego też w ostatnich dwóch latach badacze skupili się na badaniach *in vitro* właściwości katalitycznych *AID* i innych białek z rodziny *AID/APOBEC* analizując bezpośrednio ich zdolność do deaminacji 5mC do T. W wyniku tych badań rozgorzała burzliwa dyskusja nad wątpliwościami, które budzi obniżone powinowactwo centrum aktywnego *AID* wobec 5mC w porównaniu z niemodyfikowanym substratem. Wyniki niektórych badań dowodzą, że *AID* wykazuje nawet 16-krotnie niższą aktywność *in vitro* względem 5mC niż względem niemodyfikowanej cytydyny [55, 137]. Nie wyklucza to jednak możliwości istnienia białkowego kofaktora, który oddziałując z *AID* zwiększa efektywność deaminacji 5mC *in vivo*.

Badania *in vitro* wskazują także na znacząco niższą aktywność nie tylko ludzkiego białka *AID* ale również mysich *APOBEC1*, *APOBEC2* i *APOBEC3*⁷ względem 5mC w porównaniu do niemetylowanego substratu, oraz całkowity brak aktywności tych białek względem 5hmC. W badaniach tych zastosowano serię substratów zawierających 5-modyfikowaną cytydynę, tak aby każda kolejna w serii modyfikacja stanowiła większą zawadę przestrzenną w centrum aktywnym enzymu. Przetestowano odpowiednio: 5-fluoro, 5-hydrokso, 5-metylo, 5-bromo, 5-jodo, 5-hydroksymetylocytydynę. Obserwowany spadek aktywności deaminazowej wszystkich badanych białek był wprost proporcjonalny do wielkości podstawnika przy węglu C5, co według badaczy było argumentem przeciwko potencjalnemu zaangażowaniu deaminaz z rodziny *AID/APOBEC* w procesy demetylacji DNA [123].

Inna grupa badaczy pokazała natomiast, iż ludzkie białko *APOBEC3A* może w wydajny sposób deaminować 5mC w łańcuchu polinukleotydowym DNA, w przeciwieństwie do *APOBEC3G* i *AID*. W przeprowadzonych badaniach *APOBEC3A* w wydajny sposób deaminowała 5mC w komórkach *E. coli*, a jej aktywność względem 5mC była porównywalna z aktywnością względem zwykłej cytydyny. W tych samych warunkach deaminacja 5mC katalizowana przez *AID*

⁷ U myszy występuje tylko jedno białko *APOBEC3* (patrz podrozdział 1.1.1).

i APOBEC3G zachodziła w sposób niewydajny. Po wymianie segmentu wiążącego DNA w białku AID na odpowiedni segment pochodzący z APOBEC3A jej zdolność do deaminowania 5mC była wyższa. Według badaczy AID podobnie jak APOBEC3G w niespecyficzny sposób deaminuje 5mC jedynie w nieznacznym stopniu i grupa metylowa nie pasuje w pełni do kieszeni katalitycznej tego enzymu [138]. Podobną tendencję zaobserwowała inna grupa badaczy porównująca aktywność względem 5mC ludzkiego AID oraz AID z *Danio rerio* i innych ortologów AID pochodzących z ryb kostnoszkieletowych. W badaniach tych AID pochodząca z *Danio rerio* jako jedyna efektywnie deaminowała 5mC skłaniając badaczy do konkluzji, iż bezpośredni udział AID w aktywnej demetylacji jest bardziej prawdopodobny u *Danio rerio* niż u człowieka. Na podstawie modelowania badacze zasugerowali większą elastyczność centrum aktywnego AID *Danio rerio* niż człowieka [139]. Doniesienie to wspiera hipotezę, iż centrum aktywne AID mniej wydajnie deaminuje 5mC, ze względu na zawadę przestrzenną wynikającą z modyfikacji przy węglu 5, która sprawia, że 5mC nie w pełni pasuje do kieszeni katalitycznej tego enzymu. Według niniejszej hipotezy aktywność enzymatyczna AID względem danego substratu jest odwrotnie proporcjonalna do wielkości chmury elektronowej podstawnika przy węglu 5 cytydyny [124]. W rezultacie uwarunkowania steryczne postrzegane są jako znacząca bariera dla funkcjonowania deaminaz AID/APOBEC w procesie aktywnej demetylacji. Niektórzy badacze (na podstawie niskiej aktywności AID względem 5mC) sugerują nawet, iż metylacja zabezpiecza genom przed zależną od AID deaminacją [137]. Co jednak ciekawe, wydajną deaminację 5mC stwierdzono dla AID pochodzącej z *Danio rerio* oraz APOBEC3A, ale nie dla AID z innych ryb kostnoszkieletowych czy też innych białek APOBEC3. Wnioskowanie na temat potencjalnego zaangażowania białek AID/APOBEC w procesy demetylacji jedynie na podstawie wydajności reakcji deaminacji 5mC może nie być zatem całkowicie zasadne. Ponadto efektywność tej reakcji może znacząco się różnić w warunkach *in vivo*. Proponowany mechanizm niższej aktywności AID wobec 5mC w porównaniu z niemetylowanym substratem, być może zbyt pochopnie uznawany jest za argument przemawiający przeciwko udziałowi tego białka w procesach demetylacji. Z całą pewnością nie można go uznać za jednoznaczny dowód przeciwko tej tezie.

Należy też zauważyć, że żadna grupa badaczy nie pokazała do tej pory efektywności katalitycznej (wyrażonej stosunkiem stałej katalitycznej (kcat) do stałej Michaelisa (Km)) rekombinowanego enzymu AID (czy też innego z rodziny

AID/APOBEC) względem dowolnego substratu (metylowanego czy niemetylowanego). Dlatego też precyzyjne (ilościowe) określenie preferencji substratowych dla testowanych preparatów AID nie jest sprawą oczywistą.

Kolejną istotną kwestią, w rozważaniach na temat potencjalnego zaangażowania AID w procesy demetylacji, jest charakter aktywności deaminazowej AID, która w procesach immunologicznych wiąże się z aktywnością transkrypcyjną rozpoznawanych regionów genomu i specyficznym kontekstem sekwencyjnym (WRCY), w którym ulokowana jest ulegającej deaminacji cytydyna. Metylacja natomiast dotyczy motywów CpG i powszechnie wiązana jest z brakiem aktywności transkrypcyjnej zmetylowanego regionu w genomie. Nie jest jasne, w jaki sposób AID mogłaby rozpoznawać i wiązać fragmenty DNA ulegające demetylacji. W 2010 roku Bhutani i współpracownicy wykazali obecność AID związanej z regionami promotorowymi nieaktywnych transkrypcyjnie genów, zarówno w ludzkich fibroblastach jak i komórkach ES, przy czym związanie AID z tymi regionami nie skutkowało ich demetylacją [113]. Związanie AID z nieaktywnym transkrypcyjnie regionem jest zatem możliwe ale nie jest jednoznaczne z zaobserwowaniem efektów jej aktywności. Uprawdopodobnia to hipotezę, iż w takim przypadku inne czynniki np. kofaktory białkowe są niezbędne do aktywności enzymatycznej AID.

Stwierdzona metylacja DNA w obrębie aktywnych transkrypcyjnie genów nasuwa inne rozwiązanie tego problemu. Jak pokazują wyniki badań, metylacja fragmentów genów poniżej pierwszego eksonu nie wiąże się z represją transkrypcji tych genów i nie ma wpływu na procesywność polimerazy. Jedynie metylacja w obrębie pierwszego eksonu bezpośrednio blokuje inicjację transkrypcji [140]. Znaczenie i funkcja takiej wewnątrzgenowej metylacji DNA aktywnych transkrypcyjnie genów pozostaje nieznana. Niemniej jednak, najnowsze badania sugerują, że globalna hipermetylacja genomu indukowanych pluripotencjalnych mysich komórek macierzystych (ang. *induced pluripotent stem cells*, iPSC) AID -/- związana jest z metylacją regionów wewnątrzgenowych [141].

Kolejne wątpliwości budzi konieczna symetryczność procesu demetylacji na obu niciach DNA. Taka symetryczna deaminacja powodowałaby powstawanie dwuniciowych przerw w DNA podczas działania mechanizmów naprawczych (tak jak ma to miejsce w procesie CSR), co wiązałoby się z niestabilnością genomu. Trzeba jednak pamiętać, iż nasza wiedza na temat aktywności AID bazuje na jej funkcji

w procesach immunologicznych. Biochemiczne uwarunkowania aktywności AID w procesie demetylacji mogą być znacząco różne.

Choć wciąż brak jest jednoznacznych dowodów, istnieją uzasadnione przypuszczenia, że jedno lub więcej białek z rodziny AID/APOBEC może być zaangażowane w proces aktywnej demetylacji genomu oraz zjawisko pluripotencji i różnicowania komórek. Jest też możliwe, że białka AID/APOBEC komplementują się lub też wykazują w procesie demetylacji specyficzność względem różnych *loci*. Kwestia udziału AID w procesach demetylacji DNA pozostaje zatem otwarta, a wyniki badań podjętych w ramach niniejszej rozprawy są próbą włączenia się w światową dyskusję na ten temat.

1.6. Dysfunkcja AID

Mechanizmy regulacji ekspresji *AID* jak i mechanizmy warunkujące specyficzność *AID* względem *locus Ig* nie są w pełni poznane (patrz podrozdział 1.3 i 1.4). Istnieje jednak szereg dowodów na to, iż jakiegokolwiek zaburzenia w tych mechanizmach mogą prowadzić nie tylko do defektów w odpowiedzi immunologicznej, ale również do destabilizacji genomu i indukcji procesu nowotworzenia w obrębie i poza układem odpornościowym.

1.6.1. Udział *AID* w destabilizacji genomu

Zaburzona lub podwyższona ekspresja *AID* skutkuje transformacją nowotworową limfocytów B i T oraz rozwojem nowotworów w różnych organach [142, 143]. Transgeniczne myszy z konstytutywną ekspresją *AID* rozwijają chłoniaki B i T limfocytarne, jak również wiele typów nowotworów poza układem immunologicznym w tym: nowotwory płuc, wątroby, żołądka i przewodu żółciowego [144]. Co ciekawe, specyficzność organowa powstających nowotworów zależy od zastosowanego promotora [144, 145]. Również wyniki z innych systemów doświadczalnych potwierdzają te doniesienia, a heterologiczna ekspresja *AID* okazuje się prowadzić do powstawania mutacji w wielu genach poza genami immunoglobulin. Mutagenną aktywność *AID* pokazano poprzez wprowadzenie genu kodującego to białko do komórek bakteryjnych i drożdżowych [46]. Ponadto, podwyższona ekspresja *AID* w liniach komórkowych limfocytów B powoduje powstawanie mutacji we wprowadzonym do tych komórek aktywnym transkrypcyjnie transgenie kodującym białko GFP [85]. Przez sekwencjonowanie wybranych genów stwierdzono również zależną od *AID* akumulację mutacji punktowych w genomie limfocytów B zarówno w obrębie protoonkogenów: *BCL-6*, *FAS*, *PIM1*, *MYC*, *RHOH*, *PAX5*, jak i genów nieonkogennych: *B29* i *MB1*. Poziom tych mutacji korelował z aktywnością transkrypcyjną badanych genów oraz z lokalizacją motywów WRCY, co wskazuje na udział *AID* w ich generowaniu [54].

Wiązanie *AID* do rejonów genomu zlokalizowanych poza *locus Ig*. Wyniki przeprowadzonych w ostatnich latach badań genomicznych stały się katalizatorem dyskusji na temat potencjalnej aktywności *AID* wobec rejonów genomu zlokalizowanych poza *locus Ig*. Dzięki możliwości szerszego spojrzenia na genom, badacze starali się odpowiedzieć na pytanie: jak bardzo ograniczona dla *AID* jest

dostępność do innych *loci* niż *locus Ig*. Jak pokazują poniższe wyniki odpowiedź na pytanie o specyficzność działania AID, może być w rzeczywistości nie tak oczywista.

Metodą analizy genów kandydujących (ang. *candidate gene approach*) oszacowano, że AID mutuje blisko 25% wszystkich aktywnych transkrypcyjnie genów w limfocytach B w centrum namnażania limfocytów w grudkach limfatycznych. Częstość mutacji w tych genach jest jednak dużo niższa niż w genach immunoglobulin i większość z nich jest naprawiana przez mechanizmy BER i MMR. Co ciekawe, zdolność naprawy wprowadzonych przez AID mutacji jest różna w różnych miejscach genomu [146]. Wyniki te sugerują, że istotne jest nie tylko rozpoznawanie i wiązanie miejsc docelowych przez AID, ale również rozpoznawanie wprowadzonych przez AID mutacji przez maszynę naprawy DNA. W tym kontekście, wysoka zmienność genów immunoglobulin jest wynikiem nie tylko specyficznego działania AID w ich obrębie, ale też swoistego braku działania mechanizmów naprawy DNA w tych genach.

Stosując metodę ChIP-seq pokazano, iż w pierwotnej linii limfocytów B indukowanych w kierunku CSR, AID wiąże się do 5910 genów (zidentyfikowano 12 200 miejsc wiązania AID). Stopień wiązania był wyższy w regionach S, ale prawie tak samo wysoki jest w wielu innych genach, wliczając w to geny wcześniej zidentyfikowane jako potencjalnie mutowane przez AID. Większość miejsc wiązania znajdowała się w aktywnych transkrypcyjnie regionach i korelowała z miejscami wiązania kompleksu polimerazy RNA II. Co więcej, w tych samych badaniach pokazano, że miejsca wiązania dla RPA (z zaledwie kilkoma wyjątkami) ograniczają się do *locus IgH*. Wiązanie RPA do *locus IgH* było zależne od obecności AID – nie stwierdzono go w komórkach AID^{-/-} i było ono trzykrotnie zredukowane w przypadku AID z mutacją w miejscach fosforylacji: S38A lub T140A. Uzyskane wyniki pozwalają sądzić, że AID i RPA kooperatywnie wiążą się do DNA. Ponadto, pokazują one, że w pierwotnych liniach limfocytów B indukowanych w kierunku CSR, AID wiąże i mutuje wiele pozycji w genomie, mimo że preferencyjnie deaminuje regiony S *Ig* [147]. Obserwacje te stanowią istotną przesłankę wspierającą hipotezę, zgodnie z którą maszyna naprawy DNA odgrywa kluczową rolę w procesach inicjowanych przez AID.

Badania genomowe myszy przeprowadzone z wykorzystaniem macierzy ChIP-chip⁸ służących do identyfikacji białek wiążących DNA, pokazały, że nawet

⁸ ChIP-chip jest to metoda łącząca immunoprecypitację chromatyny z technologią mikromacierzy DNA.

fizjologiczny poziom AID indukuje setki DSBs w genomie komórek pochodzących z pierwotnych linii limfocytów B indukowanych w kierunku CSR. DSBs były wykrywane za pomocą białka Nbs1, które jest składnikiem kompleksu białkowego wiążącego się w odległości około 1 kbp od przerw dwuniciowych w DNA. Większość miejsc wiązania Nbs1 wykryto w regionach S *Ig*, niemniej jednak wiele z nich lokalizowano również w innych genach [148]. DSBs generowane w czasie CSR mogą prowadzić do pęknięć chromosomów i translokacji. Wyniki te ukazują skalę na jaką AID destabilizuje genom nawet w fizjologicznych warunkach. Potwierdzają tym samym hipotezę, o znaczeniu maszynarii naprawy DNA w niwelowaniu efektów działania AID w całym genomie poza *locus Ig*.

1.6.2. Udział AID w powstawaniu nowotworów

Do tej pory udowodniono udział AID w powstawaniu niektórych chłoniaków, białaczek oraz szpiczaków [149, 150]. Zaburzenia ekspresji *AID*, na różnych poziomach, zostały stwierdzone również w prawie jednej trzeciej wszystkich pierwotnych nowotworów płuc oraz w wielu innych typach nowotworów, których opis wykracza poza zakres tej pracy [59, 142]. Dane te wskazują jednak na potencjalny udział AID w powstawaniu znacznej części wszystkich nowotworów złośliwych. Poniżej krótko omówiono rolę AID w kilku wybranych procesach nowotworzenia.

95% diagnozowanych chłoniaków wywodzi się z limfocytów B, z których większość to limfocyty występujące w centrum ich namnażania w guzkach limfatycznych. Najczęstszą aberracją towarzyszącą chłoniakom limfocytarnym B-komórkowym jest translokacja chromosomowa pomiędzy *locus* immunoglobulin i protoonkogenu. W wyniku przemieszczenia fragmentu chromosomu, protoonkogen trafia pod kontrolę elementów regulatorowych *cis* genów immunoglobulin. Prowadzi to do zaburzonej, konstytutywnej ekspresji protoonkogenu i rozwoju choroby. Niektóre z translokacji są charakterystyczne dla konkretnych typów chłoniaków np.: translokacja *BCL-1/Ig* występuje w chłoniaku z komórek płaszczka (ang. *Mantle Cell Lymphoma*, MCL), translokacja *BCL-2/Ig* w chłoniaku guzkowym (ang. *Follicular Lymphoma*, FL), a translokacja *MYC/IgH* w chłoniaku Burkitta (ang. *Burkitt Lymphoma*, BL)⁹ [151, 152]. Translokacje te są wynikiem pęknięcia chromosomu w obrębie regionów zmiennych lub regionów S genów immunoglobulin. Ponieważ oba procesy: CSR i SHM

⁹ Wszystkie wymienione chłoniaki należą do chłoniaków nieziarniczych (ang. *non-Hodgkin lymphoma*, NHL).

dotyczą wyłącznie tych regionów i zachodzą jedynie w centrum namnażania limfocytów w guzłach limfatycznych, kojarzono je z rozwojem chłoniaków limfocytarnych B-komórkowych na długo przed odkryciem AID. Po jej odkryciu oraz poznaniu roli, jaką pełni w procesach różnicowania przeciwciał wykazano, że AID bezpośrednio uczestniczy w powstawaniu tych translokacji i rozwoju związanych z nimi nowotworów [134, 153, 154]. Z patogenezą chłoniaków związane są również mutacje punktowe generowane przez AID w genach supresorów nowotworzenia, takich jak: *P53*, *RB*, *P19*, *ATM* [151].

Na podstawie obserwacji, iż istnieje korelacja pomiędzy ekspresją genu AID a złymi prognozami dla przewlekłej białaczki limfatycznej B-komórkowej (ang. *B-cell chronic lymphocytic leukemia*, B-CLL) sugeruje się udział AID w patogenezie i progresji tego nowotworu [155, 156]. Co ciekawe, złe prognozy jak i zaburzona ekspresja AID w B-CLL koreluje także z brakiem hiperzmienności somatycznej regionu V genów immunoglobulin (brakiem SHM). Wskazuje to na szereg zaburzeń w mechanizmach kierowania i wiązania AID do sekwencji docelowych. Charakterystyczne dla B-CLL są również aberracje chromosomowe takie jak delecja 11q i trisomia chromosomu 12. Szczególne zainteresowanie budzi trisomia chromosomu 12, gdyż na tym chromosomie zlokalizowane jest *locus AID*. Trudno zatem jednoznacznie stwierdzić, czy zaburzona ekspresja AID jest, w tym przypadku, przyczyną czy też skutkiem procesu nowotworzenia. Wyniki badania transgenicznych myszy z konstytutywną ekspresją AID, u których obserwowano rozwój białaczek limfatycznych B-komórkowych, wskazują, że zaburzona ekspresja AID może być bezpośrednią przyczyną tego typu transformacji nowotworowej [59, 144, 157, 158].

Kolejnym nowotworem, w którego powstawanie potencjalnie zaangażowana jest AID jest szpiczak mnogi. Jest to nowotwór układu krwiotwórczego charakteryzujący się klonalną proliferacją atypowych plazmocytów. Plazmocyty są komórkami produkującymi monoklonalną immunoglobulinę powstającymi z limfocytów B jako końcowy etap ich różnicowania. Choć etiologia szpiczaków jest nieznana i w dużej mierze wieloczynnikowa, to istnieją wyniki badań wskazujące na udział AID w tym procesie. Zarówno w szpiczakowych liniach komórkowych jak i w pierwotnej linii komórek szpiczaka mnogiego (ang. *multiple myeloma*, MM) stwierdzono ekspresję AID oraz, wynikające z obecności tego białka, powstawanie DSBs i w konsekwencji destabilizację genomu [159]. Ponadto w modelu mysim wykazano istotne znaczenie

zależnej od AID aktywacji protoonkogenu *MYC* w progresji szpiczaka mnogiego z łagodniejszych form gammatopatii monoklonalnych [149].

Na uwagę zasługuje również potencjalne zaangażowanie AID w procesy nowotworzenia poprzez jej rolę w procesach demetylacji genomu. Pomimo że rola ta wciąż pozostaje kontrowersyjna, istnieją wyniki badań sugerujące związek pomiędzy zaburzeniami w regulacji potencjalnie zależnego od AID szlaku demetylacji DNA a etiologią raka okrężnicy [160]. Jest to jedyne jak dotąd doniesienie, bezpośrednio wskazujące na udział AID w procesach nowotworzenia poprzez jej udział w procesach epigenetycznych. Nie ma jednak wątpliwości, że zmiany w tak zwanym metylomie DNA są charakterystyczne dla większości (jeśli nie wszystkich) typów nowotworów [161]. Zakładając udział AID w procesach demetylacji genomu, jej zaburzona ekspresja mogłaby prowadzić do anormalnej hipometylacji DNA (m.in. regionów promotorowych protoonkogenów). Hipometylacja i w konsekwencji aktywacja takich onkogenów jak *MYC* i *RAS* jest jednym z postulowanych mechanizmów indukowania transformacji nowotworowej przez zaburzenia mechanizmów epigenetycznych [162, 163]. Ponadto globalna hipometylacja genomu jest stanem charakterystycznym dla wielu nowotworów złośliwych w tym dla: wspomnianej już B-CLL, przerzutowego pierwotnego raka wątrobowokomórkowego (ang. *metastatic hepatocellular carcinoma*, MHCC), raka prostaty i szyi [163]. Jednoznaczne potwierdzenie udziału AID w procesach demetylacji genomu otworzyłoby zatem zupełnie nowy rozdział dotyczący wpływu zaburzeń regulacji aktywności AID na procesy nowotworzenia.

1.6.3. Zaburzenia ekspresji AID powodowane przez patogeny

Zaburzona ekspresja AID może być powodowana przez niektóre patogeny, takie jak: wirusy zapalenia wątroby typu B i C, wirus Epsteina-Barr czy *Helicobacter pylori*. Patogeny te stymulują ekspresję AID w tkankach, w których normalnie gen ten pozostaje nieaktywny. Co ciekawe, każdy z tych patogenów jest potencjalnie onkogenny i związany z występowaniem charakterystycznych dla niego typów nowotworów. Sugeruje to istnienie zależnego od AID mechanizmu rozwoju nowotworów związanych ze stanem zapalnym, a zaburzona ekspresja AID staje się w ten sposób brakującym ogniwem pomiędzy atakiem patogenu a procesem nowotworzenia z nim związanym.

Wirus Epsteina-Barr (ang. *Epstein-Barr Virus*, EBV) może powodować wiele typów nowotworów w tym: raka żołądka, raka piersi, chłoniaka Burkitta oraz chorobę

Hodgkina [164]. Jak wspomniano w poprzednim podrozdziale, charakterystyczna dla chłoniaka Burkitta jest translokacja *MYC/IgH*, której powstawanie jest ściśle związane z aktywnością *AID*. Nie jest to jednak jedyny dowód na korelację infekcji EBV z ekspresją *AID*. Infekcja *in vitro* pierwotnych limfocytów B wirusem EBV skutkuje stymulacją ekspresji *AID* [165]. Co więcej, białko LMP1 wirusa EBV może indukować szlaki sygnałowe normalnie indukowane przez CD40. Ponieważ ekspresja *AID* w limfocytach B jest stymulowana przez CD40, można przypuszczać, że ekspresja *AID* indukowana infekcją EBV zachodzi w wyniku aktywacji tego samego szlaku sygnałowego [166, 167].

Podobne zaburzenia ekspresji *AID* powoduje wirus zapalenia wątroby typu C (ang. *Hepatitis C Virus*, HCV) w ludzkich hepatocytach, prowadząc w ten sposób do rozwoju raka wątroby oraz szeregu innych zaburzeń takich jak zaburzenia limfoproliferacji oligoklonalnej czy rozwój chłoniaków typu non-Hodgkin's [59]. Zaburzona ekspresja *AID* została stwierdzona zarówno w tkance wątrobowej u pacjentów zainfekowanych HCV, jak i u pacjentów z powodowanym przez HCV pierwotnym rakiem wątrobowokomórkowym (ang. *hepatocellular carcinoma*, HCC). W hepatocytach pochodzących od osób zdrowych ekspresja *AID* nie jest wykrywana [36, 168]. Znaczące jest również, że transgeniczne myszy z nadekspresją *AID* rozwijają raka wątroby o morfologicznych cechach HCC i z mutacjami w genie *P53* charakterystycznymi dla tego nowotworu [169].

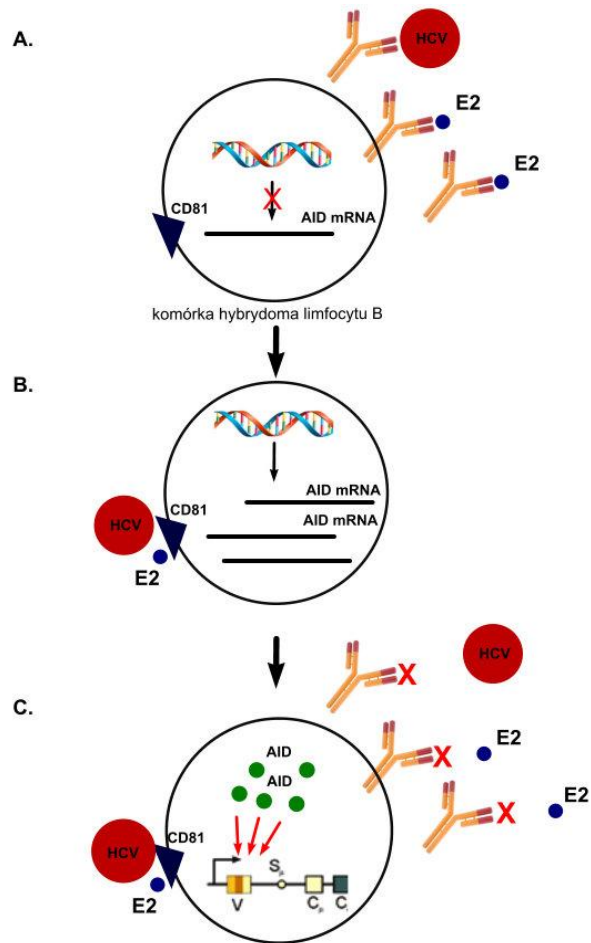
Wyniki badań Khana i współautorów sugerują, że indukcja ekspresji *AID* podczas infekcji HCV w hepatocytach odbywa się poprzez stymulację szlaku NFκB. Przypuszcza się, że szlak NFκB jest bezpośrednio indukowany przez białko C nukleokapsydu tego wirusa [170]. Indukcję tę pokazano w ludzkich liniach hepatocytów zarówno na drodze ich transfekcji wektorem zawierającym sekwencję kodującą białko C jak i w systemie replikonowym. Ponadto zaburzonej ekspresji *AID* w tych komórkach towarzyszyła akumulacja zmian nukleotydowych w genach *MYC*, *PIM1*, *P53* [36]. Szlak NFκB może być również indukowany przez cytokiny prozapalne: czynnik TNFα i interleukinę IL1β, które są wydzielane przez komórkę w wyniku kontaktu z patogenem. Inne cytokiny prozapalne: IL4 i IL13 mogą indukować zależną od infekcji ekspresję *AID* na drodze szlaku sygnałowego STAT6 [171]. Indukowaną cytokinami ekspresję *AID* wykazano zarówno w hepatocytach jak i w komórkach nabłonka dróg żółciowych [172]. W ten sposób patogeny mogą aktywować ekspresję *AID* również pośrednio, poprzez stymulację wydzielania cytokin.

Ekspresję *AID* zależną od aktywacji szlaku NFκB zaobserwowano również w komórkach błony śluzowej żołądka u osób zainfekowanych *H. pylori* [143] oraz w komórkach nowotworowych u osób z rakiem błony śluzowej żołądka [173]. Co ważne, *H. pylori* jest powszechnie uznawany za jeden z głównych czynników etiologicznych raka żołądka u ludzi. Na tej podstawie można sądzić, iż obok wirusów również bakterie są w stanie stymulować ekspresję *AID*. W przypadku *H. pylori*, obserwowane mechanizmy indukcji ekspresji *AID* były podobne jak te obserwowane dla HCV. Również w tym przypadku cytokiny prozapalne wzmacniały efekt powodowany przez sam patogen [171].

Nietypowy mechanizm potencjalnego wpływu infekcji wirusowej na odpowiedź immunologiczną gospodarza pokazała japońska grupa badaczy. Zauważyli oni, że oddziaływanie wirionu HCV (szczepu limfotropowego) lub jego białka E2 z receptorem CD81 na powierzchni komórek hybrydoma¹⁰ powodowała stymulację ekspresji *AID*. Jako następstwo działania *AID*, obserwowano akumulację mutacji punktowych w regionie zmiennym łańcucha ciężkiego immunoglobulin. W konsekwencji całego procesu infekowane komórki hybrydoma produkowały przeciwciała o obniżonym powinowactwie do antygeny w porównaniu z komórkami nieinfekowanymi (Rysunek 1.6.1) [174, 175]. Jest to potencjalnie nowy, nieznany dotąd mechanizm, na drodze którego wirus może modulować odpowiedź immunologiczną gospodarza przez wpływ na specyficzność produkowanych przeciwciał. Choć istnienie takiego mechanizmu musi znaleźć potwierdzenie jeszcze w innych systemach badawczych, wyniki te są przykładem zaskakujących konsekwencji zaburzonej obecnością patogenu ekspresji *AID*.

W świetle omówionych wyników badań, *AID* staje się niezwykle obiecującym celem terapii nie tylko schorzeń immunologicznych, ale również nowotworowych i wirusowych. Niestety, jak do tej pory nie jest znany żaden związek specyficznie modulujący aktywność *AID* (inhibitor czy też aktywator). Szczegółowe zbadanie mechanizmów molekularnych leżących u podłoża aktywności biologicznej tego białka otworzy zatem drogę do opracowania związków będących potencjalnym źródłem nowych leków.

¹⁰ Komórki hybrydoma to komórki powstałe w wyniku fuzji komórek szpiczaka z klonem limfocytów B produkującym określoną immunoglobulinę.



Rysunek 1.6.1. Schemat ilustrujący potencjalny mechanizm modulacji odpowiedzi immunologicznej gospodarza przez HCV. A) Komórka hybrydoma produkująca przeciwciała specyficzne wobec białka E2 wirusa HCV. B) Oddziaływanie wirionu HCV lub jego białka E2 z receptorem CD81 na powierzchni komórek linii hybrydoma stymuluje ekspresję *AID*. C) *AID* inicjuje proces SHM, w efekcie którego komórki hybrydoma produkują przeciwciała o obniżonym powinowactwie do patogenu. Brak oddziaływania przeciwciał z antygenem zaznaczono na rysunku czerwonym znakiem „X”. Na podstawie [174, 175].

2. CEL PRACY

Mimo że od odkrycia AID upłynęło już 16 lat, białko to jest stosunkowo słabo poznane i nadal nie znaleziono odpowiedzi na wiele pytań dotyczących jego aktywności biologicznej. Najwięcej kontrowersji budzi potencjalna aktywność tego białka wobec modyfikowanych cytydyn, a co za tym idzie jego funkcje poza układem immunologicznym. Proponowany jest udział AID w procesach aktywnej demetylacji genomu zarówno poprzez deaminację 5-metylocytydyny do tymidyny, jak i deaminację 5-hydroksymetylocytydyny do 5-hydroksymetylourydyny. Szereg doniesień (pochodzących głównie z systemów *in vivo*) potwierdza aktywność AID wobec modyfikowanych substratów. Jednocześnie, równie licznie, pojawiają się dane, które przeczą takiej możliwości. Głównym celem badań podjętych w ramach niniejszej pracy doktorskiej stała się zatem charakterystyka mechanizmów molekularnych leżących u podłoża aktywności biologicznej ludzkiej AID (ang. *human AID*, hAID), a w szczególności jej aktywności enzymatycznej wobec cytydyny, 5-metylocytydyny i 5-hydroksymetylourydyny *in vitro*.

Jedną z przyczyn uniemożliwiających udzielenie jednoznacznej odpowiedzi na zasadnicze pytania dotyczące aktywności hAID był brak wydajnego systemu do produkcji tego białka. Dlatego w pierwszym etapie badań postawiono sobie za cel stworzenie systemu produkcji hAID o pełnej aktywności enzymatycznej. Uzyskane tą drogą białka miały posłużyć do szczegółowej charakterystyki właściwości enzymatycznych hAID i jej mutantów w warunkach *in vitro*. Spodziewano się, że analizy te pozwolą zidentyfikować czynniki strukturalne warunkujące aktywność hAID wobec modyfikowanych i niemodyfikowanych substratów.

Osiągnięcie tak zarysowanego celu wymagało realizacji następujących zadań szczegółowych:

- 1) stworzenia wydajnego systemu produkcji hAID o pełnej aktywności enzymatycznej;
- 2) opracowania metod testowania aktywności deaminazowej hAID wobec modyfikowanych i niemodyfikowanych substratów w warunkach *in vitro*;
- 3) otrzymania mutantów hAID wykazujących zróżnicowaną aktywność *in vivo*;
- 4) przeprowadzenia szczegółowej charakterystyki aktywności enzymatycznej hAID typu dzikiego oraz wyselekcjonowanych mutantów wobec modyfikowanych i niemodyfikowanych substratów;

- 5) identyfikacji uwarunkowań strukturalnych aktywności hAID wobec cytydyny i 5-metylo/hydroksymetylocytydyny.

3. MATERIAŁY I METODY

3.1. Wykaz materiałów stosowanych w pracy

3.1.1. Podstawowe odczynniki, złoże, enzymy oraz zestawy gotowych odczynników

Tabela 3.1.1. Podstawowe odczynniki stosowane w pracy¹¹.

Odczynnik	Producent
NucAway™ Spin Columns	Ambion
agar, 40% akrylamid/bis-akrylamid (29:1), ampicylina, chlorek sodu, chlorek wapnia (1 M roztwór), chlorek potasu, chloramfenikol, dodecylosiarczan sodu (SDS), glicerol, glicyna, kwas etylenodiamino-tetraoctowy (EDTA), kwaśny fosforan sodu (NaH ₂ PO ₄), mocznik, ortofosforan dwusodowy (Na ₂ HPO ₄), pepton, siarczan magnezu, surowicza albumina wołowa (BSA), TBE, TBST z Tween-em 20 (TBSTT), Tris, wyciąg drożdżowy	Bioshop
5-bromo-4-chloro-3-indolo-D-galaktozyd (X-Gal), bufor do ligacji (10x Ligation Buffer), bufor reakcyjny dla kinazy polinukleotydowej T4 (10x PNK A), bufor Taq do reakcji PCR, chlorek magnezu, izopropylotioegalaktozyd (IPTG), odczynnik Bradford, trifosforany deoksyrybonukleotydów (dATP, dCTP, dGTP, dTTP 100 mM, dNTPs)	Fermentas
γ [³² P]dATP (3000 Ci/mmol)	ICN
β -merkaptioetanol	MERCK
filtry 10 kDa do zagęszczania białek, filtry 0,22 μ m do sterylizacji roztworów, tetrahydryourydyna (THU)	MERCK Millipore
bufor AP (20x)	Novagen
membrana PVDF 0,45 μ m	Thermo Scientific
etanol, kwas octowy, metanol, wodorotlenek sodu	Polskie Odczynniki Chemiczne
Agaroz	PRONA
inhibitor proteaz Complete Mini EDTA-free Coctail	Roche Diagnostics
Serva Blue R-250	SERVA
Formamid	Applied Biosystems
1,10-fenantrolina, błękit bromofenolowy, bromek etydyny (10 mg/ml), chloroform, chlorowodorek guanidyny, dimetyloformamid (DMF), dimetylosulfotlenek (DMSO), ditiotreitrol (DTT, 1 M roztwór), fenol, imidazol, kanamycyna, L-arabinoza, lizozym, nadsiarczan amonu (APS), <i>n</i> -Dodecyl β -D-maltozyd (DDM), octan amonu, octan sodu, siarczan niklu, spektynomycyna, tetrametyloetylenodiamina (TEMED), tris(2-karboksyetylo)fosfina (TCEP), Triton X-100, Tween 20, wodorofosforan potasu (K ₂ HPO ₄), zredukowany glutation	Sigma-Aldrich

¹¹ Jeśli nie zostało określone inaczej, podane stężenie procentowe danej substancji oznacza jej stężenie wagowo/objętościowe (w/v). W przypadku, gdy nie jest sprecyzowany rodzaj rozpuszczalnika, należy rozumieć, że jest nim H₂O. v/v – stężenie objętościowo/objętościowe.

Tabela 3.1.2. Złoże stosowane w pracy.

Złoże	Producent
złoże glutationowe (Gluthatione Sepharose 4 Fast Flow)	GE Healthcare
złoże heparynowe (Heparin Sepharose 6 Fast Flow)	
złoże Superdex75 16/60 HiLoad	
złoże niklowe (Ni-NTA Agarose)	Qiagen
DEAE celuloza	Sigma-Aldrich

Tabela 3.1.3. Enzymy stosowane w pracy.

Enzym	Producent
ligaza DNA T4 (5 U/μl), kinaza polinukleotydowa T4 (10 U/μl), polimeraza DNA Taq (5 U/μl), RNaza A (10 μg/μl), glikozylaza uracylo-DNA (UDG, 1 U/μl), BamH I (10 U/μl), EcoRI (10 U/μl), DpnI (10 U/μl), Fast Digest® BamH I, Fast Digest® EcoRI, Fast Digest® EcoRV	Fermentas
benzonase® nukleaza (25 U/μl)	Novagen
ludzka jednofunkcyjna glikozylaza uracylo-DNA selektywna wobec ssDNA (hSMUG1, 5 U/μl)	New England Biolabs
polimeraza DNA Pfu (3 U/μl)	Promega
termostabilna glikozylaza tymidyno-DNA (TDG, 0,5 U/μl)	Trevigene

Tabela 3.1.4. Zestawy gotowych odczynników stosowane w pracy.

Zestaw gotowych odczynników	Producent
zestaw do sekwencjonowania DNA: BigDye® Terminator V1.1 cycle sequencing kit	Applied Biosystems
zestaw do kolorymetrycznej detekcji obecności fosfatazy alkalicznej (w procedurze <i>western blot</i>): AP Detection Reagent Kit	Novagen
zestaw do klonowania produktów PCR do plazmidu pGem®-T Easy: pGem®-T Easy Vector System	Promega
zestaw do oczyszczania DNA: QIAquick PCR Purification Kit, zestaw do oczyszczania plazmidów w średniej skali (do 100 μg): Plasmid Midi Kit	Qiagen
zestaw do badania globalnego poziomu metylacji genomowego DNA: 5-mC DNA ELISA Kit	Zymo Research
zestaw do amplifikacji DNA zawierającego uracyl: KAPA HiFi HotStart Uracil+ ReadyMix	KAPA Biosystems
zestaw do cięcia proteolitycznego białek: Enterokinase Cleavage Capture Kit	Novagen

3.1.2. Oligodeoksyrybonukleotydy

Oligonukleotydy DNA stosowane podczas klonowania genu AID i otrzymywania jej mutantów były syntetyzowane przez Genomed S.A. (Warszawa). Oligonukleotydy DNA stosowane w badaniach aktywności AID były syntetyzowane przez IBA GmbH (Göttingen, Niemcy).

Tabela 3.1.5. Zestawienie oligodeoksyrybonukleotydów wykorzystywanych w pracy.

Nazwa	Sekwencja 5'-3'
Oligodeoksyrybonukleotydy stosowane podczas klonowania genu AID i otrzymywania jej mutantów	
F_03_aid	GACGACGACGACAAGATGGACAGCCTCTTGATGAAC
F_04_aid	ATGGATCCACCATCATCATCATGACGACGACGACAAGATGGA
R_03_aid	ATGAATTCCTATCAAAGTCCCAAAGTACGA
Fwew_aid	ATGTGGCCGACTTTCTGCGA
Rwew_aid	TGAAGATCCTCAGACTGAGGT
mutN51A_F	CTTTGGTTATCTTCGCGCTAAGAACGGCTGCC
mutN51A_R	GAAACCAATAGAAGCGCGATTCTTGCCGACGG
mutR50A_F	GACTTTGGTTATCTTGCCAATAAGAACGGCTG
mutR50A_R	CAGCCGTTCTTATTGGCAAGATAACCAAAGTC
mut Δ189_R	ATGAATTCCTATCAGTCATCAACCTCATACAGGGG
Oligodeoksyrybonukleotydy stosowane w badaniach aktywności AID	
Oligo_C	<u>GGATTGGTTGGTTATTTGTT</u> TAAGGAAGGTGGATTA AAAAGCTTAATAAGG TGATGGAAGTTATGTTTGGTAGATTGATGG
Oligo_U	<u>GGATTGGTTGGTTATTTGTT</u> TAAGGAAGGTGGATTA AAAAGUTTAATAAGG TGATGGAAGTTATGTTTGGTAGATTGATGG
Oligo_5mC	<u>GGATTGGTTGGTTATTTGTT</u> TAAGGAAGGTGGATTA AAAAG [5-metylo-dC]TTAATAAGGTGATGGAAGTTATGTTTGGTAGATTGATGG
Oligo_T	<u>GGATTGGTTGGTTATTTGTT</u> TAAGGAAGGTGGATTA AAAAGTTTAATAAGG TGATGGAAGTTATGTTTGGTAGATTGATGG
Oligo_5hmC	<u>GGATTGGTTGGTTATTTGTT</u> TAAGGAAGGTGGATTA AAAAG [5-hydroksymetylo-dC]TTAATAAGGTGATGGAAGTTATGTTTGGT AGATTGATGG

Nazwa	Sekwencja 5'-3'
Oligo_5hmU	<u>GGATTGGTTGGTTATTTGTTTAAGGAAGGTGGATTA</u> AAAAG [5-hydroksymetylo-dU]TTAATAAGGTGATGGAAGTTATGTTTGGTA <u>GATTGATGG</u>
Oligo_C- komplement	<u>CCATCAATCTACCAAACATA</u> ACTTCCATCACCTTATTAAGCTTTTAATCCA CCTTCCTTAAACAATAACCAACCAATCC
NNCN_1	<u>GGATTGGTTGGTTATTTGTTAGCAGACGACCCACGAGCTGACCTCAC</u> ACCAGCCGACTTCCCACTAGCGGACATTCGCACATATGTTTGGTAGAT <u>TGATGG</u>
NNCN_2	<u>GGATTGGTTGGTTATTTGTTATCAGTCGAACCCTCGACCTGTCCTGCTCT</u> CCATCCGCCTTACCGCCATCGGTCATACGCTCATATGTTTGGTAGATT <u>GATGG</u>
NNCN_3	<u>GGATTGGTTGGTTATTTGTTAACAGGCGTCCACCCGAACTGCCATCCTC</u> GCAACCAGGCTGGCATGCACCCTAACGTCCGCCATATGTTTGGTAGAT <u>TGATGG</u>
NNCN_4	<u>GGATTGGTTGGTTATTTGTTATCTGGCCTACTCCCCTTCTGCCGTGCCCG</u> CGACCGCGCTTGCGCGCCTACAGTCTTCCCCTCTTATGTTTGGTAGATTG <u>ATGG</u>
pcrNNCN_F	GGATTGGTTGGTTATTTGTT
pcrNNCN_R	CCATCAATCTACCAAACATA
R_pGem	TCAAGCTATGCATCCAACGCG
R2_pGem	AGTTAGCTCACTCATTAGGCA

Kolorami oznaczono: (i) pomarańczowym – sekwencje kodujące miejsce rozpoznawane przez enterokinazę; (ii) czerwonym – kodony STOP; (iii) zielonym – sekwencje rozpoznawane przez enzymy restrykcyjne; (iv) niebieskim – sekwencję kodującą znacznik histydynowy (HisTag); (v) fioletowym – nukleotydy wprowadzające żądaną mutację do sekwencji aminokwasowej AID. W oligonukleotydach DNA stosowanych w badaniach aktywności AID podkreśleniem zaznaczono miejsca wiązania starterów: pcrNNCN_F i pcrNNCN_R. Czcionką pogrubioną zaznaczono nukleotydy potencjalnie deaminowane przez AID. Pogrubieniem i kursywą zaznaczono nukleotydy będące potencjalnym produktem deaminacji.

3.1.3. Plazmidy

- pOTB7 firmy GeneCopoeia™ z wklonowanym cDNA genu ludzkiej AID – wektor replikacyjny umożliwiający klonowanie cDNA AID i jego namnażanie w komórkach bakteryjnych, zawiera gen oporności na chloramfenikol;

- pGEM[®]-T Easy firmy PROMEGA – wektor replikacyjny do klonowania fragmentów DNA i namnażania ich w komórkach bakteryjnych, zawiera gen oporności na ampicylinę, umożliwia klonowanie produktów PCR bez użycia enzymów restrykcyjnych – z wykorzystaniem niesparowanych reszt adenozynowych na końcach 3' (ang. *TA cloning*), test α -komplementacji umożliwia selekcję komórek zawierających wektor z insertem;
- pGEX-6p-3 firmy GE Healthcare – wektor ekspresyjny, zawiera gen oporności na ampicylinę, promotor *tac*, umożliwia otrzymanie białka fuzyjnego z białkiem GST i miejscem rozpoznawanym przez proteazę PreScission;
- pGEX-4T-1 firmy GE Healthcare – wektor ekspresyjny, zawiera gen oporności na ampicylinę, promotor *tac*, umożliwia otrzymanie białka fuzyjnego z białkiem GST i miejscem rozpoznawanym przez trombinę;
- pCDFDuet-1 firmy NOVAGEN z wklonowaną sekwencją kodującą ludzką AID – konstrukt otrzymany dzięki uprzejmości dr Niny Papavasiliou z The Rockefeller University (Nowy Jork, USA), wektor ekspresyjny, zawiera gen oporności na spektynomycynę, promotor T7, umożliwia otrzymanie białka fuzyjnego AID ze znacznikiem histydynowym (HisTag) i miejscem rozpoznawanym przez proteazę 3C.

3.1.4. Szczepy bakteryjne

Szczepy bakteryjne stosowane do namnażania i selekcji klonów:

- *E. coli* DH5 α (producent: Gibco BRL)
- *E. coli* One shot[®] TOP10 (producent: Life Technologies)

Szczepy bakteryjne stosowane do produkcji białka:

- *E. coli* BL21(DE3)pLysS (producent: Novagen)
- *E. coli* Origami(DE3)pLysS (producent: Novagen)
- *E. coli* BL21-AL[™] One Shot[®] (producent: Invitrogen)
- *E. coli* Rosetta(DE3)pLysS (producent: Novagen)

Tabela 3.1.6. Zestawienie wybranych cech szczepów bakteryjnych stosowanych w produkcji białka.

szczep <i>E.coli</i>	oporność na antybiotyki	promotor genu polimerazy RNA T7	gen kodujący lizozym	inne cechy
BL21(DE3)pLysS	chloramfenikol	<i>lac</i>	obecny	brak proteaz <i>lon</i> i <i>ompT</i>
Origami(DE3)pLysS	chloramfenikol, kanamycyna, tetracyklina	<i>lac</i>	obecny	mutacje w genach kodujących reduktazę tioredoksynową i glutationową
BL21-AL™ OneShot®	-	<i>araBAD</i>	brak	-
Rosetta(DE3)pLysS	chloramfenikol	<i>lac</i>	obecny	geny kodujące tRNA dla sześciu kodonów o niskiej częstotliwości występowania w <i>E.coli</i>

3.1.5. Przeciwciała

- kozie poliklonalne przeciwciała przeciwko ludzkiej AID (ang. *anti-AID polyclonal antibody produced in goat*) 0,5 mg/ml, stosowane rozcieńczenie 1:5000, producent: Sigma-Aldrich
- królicze poliklonalne przeciwciała przeciwko kozim IgG sprzężone z fosfatazą alkaliczną (ang. *rabbit polyclonal Ab to goat IgG (AP)*) 1 mg/ml, stosowane rozcieńczenie 1:5000, producent: ABCAM

3.1.6. Roztwory i bufor

Roztwory do izolacji plazmidów w małej skali

Roztwór ALM1

- 50 mM dekstroza
- 10 mM EDTA
- 25 mM Tris

Roztwór ALM2 (świeżo przygotowany)

- 1% SDS
- 0,2 M NaOH

10x stężony bufor TBE do rozdzielania elektroforetycznego kwasów nukleinowych w żelach agarozowych

- 0,9 M Tris
- 0,9 M kwas borowy

- 25 mM EDTA

Gotową mieszaninę składników firmy Bioshop rozpuszczano w 1 l H₂O, uzyskując bufor 10x stężony. Do rozdzielania elektroforetycznego w żelach agarozowych stosowano bufor 0,5x stężony. Do rozdzielania elektroforetycznego w żelach poliakrylamidowych stosowano bufor 1x stężony.

10x stężony bufor do elektroforezy białek w warunkach denaturujących (10x R)

- 0,25 M Tris-HCl, pH = 8.3
- 1,92 M glicyna
- 1% SDS

Bufor R stosowano w stężeniu końcowym 1x.

Roztwory do wybarwiania białek rozdzielonych metodą elektroforezy w żelu poliakryloamidowym

Roztwór barwiący

- 50%, v/v metanol
- 10%, v/v kwas octowy
- 0,125% Serva Blue R-250

Pierwszy roztwór odbarwiający

- 5%, v/v metanol
- 12,5%, v/v kwas octowy

Roztwory i bufor stosowane w technice *western blot*

Bufor do elektrotransferu

- 1x bufor R
- 20%, v/v metanol

Bufor TBSTT

- 25 mM Tris
- 137 mM NaCl
- 2,7 mM KCl
- 0,05% Tween-20

Gotową mieszaninę składników firmy Bioshop rozpuszczano w 1 l H₂O, uzyskując bufor 1x stężony.

Roztwór blokujący

- 1% roztwór BSA w buforze TBSTT

Roztwór przeciwciał (I- lub II-rzędowych)

- rozcieńczenie 1:5000 w roztworze blokującym (0,5 µl stężonego przeciwciała / 2,5 ml roztworu blokującego)

Roztwór do detekcji kolorymetrycznej

Przygotowywany zgodnie z zaleceniem producenta zestawu do kolorymetrycznej detekcji obecności fosfatazy alkalicznej: AP Detection Reagent Kit (Novagen).

- 15 ml jednokrotnie stężonego buforu AP
- 60 µl NBT
- 60 µl BCIP

Bufory do analizy rozpuszczalności białek**Bufor A:**

- 50 mM Na₂HPO₄, pH = 8
- 300 mM NaCl

Bufor B:

- 100 mM Na₂HPO₄, pH = 8
- 10 mM Tris-HCl
- 8 M mocznik

Bufory do oczyszczania białka na kolumnie wypełnionej złożem Ni-NTA**Bufor I:**

- 50 mM Na₂HPO₄, pH = 8
- 50 mM NaCl

Bufor II:

- 50 mM Na₂HPO₄, pH = 7
- 300 mM NaCl

Bufor III 2x stężony:

- 100 mM Na₂HPO₄, pH = 6
- 600 mM NaCl
- 20% glicerol (w/v)

Bufor III stosowano w stężeniu końcowym 1x.

Bufory do regeneracji złoża Ni-NTA

Bufor 1 (wymywający pozostałe na złożu białka):

- 1 M NaCl
- 1x bufor III

Bufor 2 (wymywający jony Ni²⁺ ze złoża):

- 20 mM Tris-HCl, pH = 7,9
- 0,5 M NaCl
- 100 mM EDTA

Bufor 3 (ładujący złożę jonami Ni²⁺):

- 50 mM NiSO₄

Bufory do sączenia molekularnego

Bufor SE1:

- 50 mM Tris-HCl, pH = 8
- 25 mM NaCl
- 3% glicerol
- 1 mM DTT

Bufor SE2:

- 50 mM Tris-HCl, pH = 8
- 200 mM NaCl
- 3% glicerol
- 0,01% DDM
- 2mM TCEP

Bufory do oczyszczania białek na kolumnie wypełnionej złożem glutationowym

Bufory przygotowano zgodnie z zaleceniami producenta złoża glutationowego (GE Healthcare). Skład buforów podano poniżej.

Bufor PBS 10x stężony:

- 100 mM Na₂HPO₄
- 18 mM KH₂PO₄
- 1,4 M NaCl
- 27 mM KCl

Bufor PBS stosowano w stężeniu końcowym 1x.

Bufor do elucji:

- 50 mM Tris-HCl, pH = 8
- 10 mM glutation zredukowany
- 1 mM DTT
- 3% glicerol

Bufory do elucji białek z kolumny wypełnionej złożem heparynowym (H1 – H5)

- 50 mM Tris-HCl, pH = 8
- 3% glicerol
- 50 mM, 130 mM, 250 mM, 600 mM lub 2 M NaCl (odpowiednio w buforach elucyjnych: H1, H2, H3, H4, H5)

Bufor reakcyjny dla enterokinazy

- 20 mM Tris-HCl, pH = 7,4
- 50 mM NaCl
- 2 mM CaCl₂

Bufor do przechowywania białka AID

- 50 mM Tris-HCl, pH = 8
- 1 mM DTT
- 50% glicerol

10x stężony bufor do badania aktywności hAID

- 500 mM Tris-HCl, pH = 8
- 10 mM DTT

Bufor reakcyjny dla białka AID stosowano w stężeniu końcowym 1x.

Inne:

roztwór chloramfenikolu w alkoholu etylowym (34 mg/ml)

roztwór X-Gal (0,5 mg/ml) w dimetyloformamidzie (DMF)

wodny roztwór ampicyliny (100 mg/ml)

wodny roztwór APS (10%)

wodny roztwór CH₃COONH₄ (7,5 M)

wodny roztwór EDTA pH = 8 (25 mM, 125 mM, 250 mM lub 0,5 M)

- wodny roztwór glicerolu (80%)
- wodny roztwór imidazolu (1 M)
- wodny roztwór IPTG (1 M lub 0,1 M)
- wodny roztwór kanamycyny (30 mg/ml)
- wodny roztwór L-arabinozy (20%)
- wodny roztwór lizozymu (50 mg/ml)
- wodny roztwór NaCl (4 M)
- wodny roztwór Na₂HPO₄, pH = 6, 7, 7,5 lub 8 (0,5 M)¹²
- wodny roztwór MgSO₄ (1 M)
- wodny roztwór SDS (10%)
- wodny roztwór spektynomycyny (50 mg/ml)
- wodny roztwór ZnCl₂ (1 mM, 100 mM)

3.1.7. Pożywki

Pożywka stała LB z ampicyliną lub spektynomycyną

- 2 g peptonu
- 1,25 g wyciągu drożdżowego
- 1,25 g NaCl
- 3,75 g agaru
- dopełniano H₂O do 250 ml i sterylizowano w autoklawie
- po schłodzeniu do temperatury 55°C do otrzymanego roztworu dodawano 500 µl roztworu ampicyliny o stężeniu 0,1 g/ml (stężenie końcowe 200 µg/ml) lub 250 µl roztworu spektynomycyny o stężeniu 50 mg/ml (stężenie końcowe 50 µg/ml) po czym na stole z laminarnym przepływem powietrza wylewano pożywkę na sterylne szalki.

Pożywka płynna LB

- 10 g peptonu
- 5 g wyciągu drożdżowego
- 5 g NaCl
- dopełniano H₂O do 1 l i sterylizowano w autoklawie

¹² otrzymywano poprzez zmieszanie wodnego roztworu Na₂HPO₄ (0,5 M) i wodnego roztworu NaH₂PO₄ (0,5 M) w odpowiednich dla danej wartości pH proporcjach

Pożywka płynna LB z ampicylina

- do 1 l płynnej pożywki LB (schłodzona po sterylizacji) dodawano 2 ml roztworu ampicyliny o stężeniu 0,1 g/ml (stężenie końcowe 200 µg/ml)

Pożywka płynna LB z ampicylina i chloramfenikolem

- do 1 l płynnej pożywki LB (schłodzonej po sterylizacji) dodawano 2 ml roztworu ampicyliny o stężeniu 0,1 g/ml (stężenie końcowe 200 µg/ml) oraz 1 ml roztworu chloramfenikolu o stężeniu 34 mg/ml (stężenie końcowe 34 µg/ml)

Pożywka płynna LB z ampicylina, chloramfenikolem i kanamycyna

- do 1 l płynnej pożywki LB (schłodzona po sterylizacji) dodawano 2 ml roztworu ampicyliny o stężeniu 0,1 g/ml (stężenie końcowe 200 µg/ml), 1 ml roztworu chloramfenikolu o stężeniu 34 mg/ml (stężenie końcowe 34 µg/ml) oraz 1 ml roztworu kanamycyny o stężeniu 30 mg/ml (stężenie końcowe 30 µg/ml).

Pożywka płynna LB z ampicylina, chloramfenikolem i cynkiem

- do 1 l płynnej pożywki LB (schłodzonej po sterylizacji) dodawano 2 ml roztworu ampicyliny o stężeniu 0,1 g/ml (stężenie końcowe 200 µg/ml), 1 ml roztworu chloramfenikolu o stężeniu 34 mg/ml (stężenie końcowe 34 µg/ml) oraz 0,6 ml roztworu chlorku cynku ($ZnCl_2$) o stężeniu 100 mM (stężenie końcowe 0,06 mM).

Pożywka płynna LB ze spektynomycyna i chloramfenikolem

- do 1 l płynnej pożywki LB (schłodzona po sterylizacji) dodawano 1 ml roztworu spektynomycyny o stężeniu 50 mg/ml (stężenie końcowe 50 µg/ml) oraz 1 ml roztworu chloramfenikolu o stężeniu 34 mg/ml (stężenie końcowe 34 µg/ml)

Płynna pożywka SOC

Gotowa pożywka dostarczona wraz z komórkami kompetentnymi firmy Novagen.

3.1.8. Żele

Żel agarozowy 1% do rozdziału elektroforetycznego kwasów nukleinowych w warunkach natywnych

- 0,45 g agarozy
- 45 ml 0,5x stężonego buforu TBE

Rozpuszczano przez podgrzanie w kuchenke mikrofalowej, po czym schładzano, dodawano bromku etydyny do stężenia końcowego 0,5 µg/ml i wylewano do aparatu do elektroforezy.

Żel poliakryloamidowy 15% do rozdziału elektroforetycznego kwasów nukleinowych w warunkach denaturujących

- 18,75 ml 40% akrylamid/N,N'-metylenobis-akrylamid
- 5 ml buforu 10x TBE
- 21 g mocznika
- dopełniano H₂O do 50 ml
- 300 µl 10% APS
- 32 µl TEMED

Wylewano między szyby i pozostawiano do polimeryzacji. Podano objętości odczynników na żel o wymiarach 13/19,5 cm.

Żele poliakryloamidowe do rozdziału elektroforetycznego białek**Żel rozdzielający 15%**

- 4,5 ml 40% akrylamid/N,N'-metylenobis-akrylamid
- 4,4 ml 1 M buforu Tris-HCl, pH = 8,8
- 120 µl 10% SDS
- 120 µl 10% APS
- 6 µl TEMED
- 2,98 ml H₂O

Żel rozdzielający 12%

- 3,6 ml 40% akrylamid/N,N'-metylenobis-akrylamid
- 4,4 ml 1 M buforu Tris-HCl, pH = 8,8
- 120 µl 10% SDS
- 120 µl 10% APS
- 6 µl TEMED

- 3,88 ml H₂O

Żel rozdzielający 12% lub 15% wylewano między szyby i spryskiwano 1% SDS, pozostawiano do polimeryzacji.

Żel zagęszczający 5%

- 625 µl 40% akrylamid/N,N'-metylenobis-akrylamid
- 630 µl 1 M buforu Tris-HCl, pH = 6,8
- 50 µl 10% SDS
- 50 µl 10% APS
- 5 µl TEMED
- 3,6 ml H₂O

Wylewano między szyby powyżej wcześniej spolimeryzowanego żelu rozdzielającego, pozostawiano do polimeryzacji.

Podano objętości odczynników na dwa żele o wymiarach żelu rozdzielającego 7,5/6,5 cm.

3.1.9. Bufory obciążające i wzorce mas

6x stężony bufor obciążający (6x BJ) do rozdzielania elektroforetycznego kwasów nukleinowych w żelach agarozowych, w warunkach natywnych

- 10 mM Tris-HCl, pH = 7,5
- 50 mM EDTA
- 10% Ficoll 400
- 0,25% błękit bromofenolowy
- 0,25% cyjanian ksylenu

2x stężony bufor obciążający (2x BB) do rozdzielania elektroforetycznego kwasów nukleinowych w żelach poliakryloamidowych, w warunkach denaturujących

- 0,25% błękit bromofenolowy
- 7 M mocznik

4x stężony bufor (4x SB) do rozdzielania elektroforetycznego białek w żelu poliakryloamidowym, w warunkach denaturujących

- 250 mM Tris-HCl, pH = 6,8
- 20% v/v β-merkaptoetanol
- 8% SDS

- 0,15% błękit bromofenolowy
- 40% glicerol

wzorzec mas DNA do rozdziału elektroforetycznego w żelach agarozowych w warunkach natywnych

- DirectLoad™ Wide Range DNA Marker (producent: SIGMA)

wzorzec mas białek do rozdziału elektroforetycznego w żelach poliakryloamidowych w warunkach denaturujących

- Novex® Sharp Pre-Stained Protein Standards (producent: INVITROGEN)
(wzorzec stosowany również podczas procedury *western blot*)

3.2. Wykaz metod stosowanych w pracy

3.2.1. Metody stosowane w pracy z kwasami nukleinowymi

3.2.1.1. Łańcuchowa reakcja polimerazy (PCR)

Amplifikację DNA metodą PCR stosowano w celu powielenia fragmentów DNA, wprowadzenia do nich miejsc restrykcyjnych lub też innych sekwencji (np. kodujących znacznik histydynowy). Stosowano polimerazę Taq (producent: Fermentas) lub też polimerazę Pfu (producent: Promega) o wysokiej dokładności kopiowania. Skład oraz warunki typowych reakcji PCR odpowiednio dla obu polimeraz przedstawiono w Tabelach 3.2.1. – 3.2.4. Reakcje prowadzono w termocyklerze firmy Biometra.

Tabela 3.2.1. Skład mieszaniny reakcyjnej - PCR z wykorzystaniem polimerazy Taq.

odczynnik (stężenie początkowe)	ilość (stężenie końcowe)
matryca DNA	1-100 ng
starter homologiczny do końca 5' powielanego fragmentu	15 pmol
starter komplementarny do końca 3' powielanego fragmentu	15 pmol
bufor reakcyjny 10x (dostarczony wraz z polimerazą Taq przez firmę Fermentas)	5 µl
MgCl ₂ (25 mM)	6 µl (3 mM)
mix dNTPs (2,5 mM każdy)	4 µl (0,2 mM)
polimeraza DNA Taq (5 U/µl)	0,5 µl (2,5 U)
H ₂ O	do 50 µl

Tabela 3.2.2. Warunki PCR z wykorzystaniem polimerazy Taq.

numer etapu	etap	czas trwania	temperatura	liczba cykli
1	denaturacja	2 min	94°C	-
2	denaturacja	1 min	94°C	30
	hybrydyzacja	30 s	(T _M)*	
	elongacja	1 min	72°C	
3	elongacja	7 min	72°C	-

* zależna od temperatury topnienia użytych starterów - T_M obliczono wg wzoru

$$T_M = 2(A+T) + 4(G+C) - 5 \text{ [}^\circ\text{C]}$$

Tabela 3.2.3. Skład mieszaniny reakcyjnej – PCR z wykorzystaniem polimerazy Pfu.

odczynnik (stężenie początkowe)	ilość (stężenie końcowe)
matryca DNA	1-100 ng
starter homologiczny do końca 5' powielanego fragmentu	15 pmol
starter komplementarny do końca 3' powielanego fragmentu	15 pmol
bufor reakcyjny 10x (dostarczony wraz z polimerazą Pfu przez firmę Promega)	5 µl
mix dNTPs (10 mM każdy)	1 µl (0,2 mM)
polimeraza DNA Pfu (3 U/µl)	0,5 µl (1,5 U)
H ₂ O	do 50 µl

Tabela 3.2.4. Warunki PCR z wykorzystaniem polimerazy Pfu.

numer etapu	etap	czas trwania	temperatura	liczba cykli
1	denaturacja	2 min	95°C	-
2	denaturacja	1 min	95°C	30
	hybrydyzacja	30 s	(T _M)*	
	elongacja	2,5 min	72°C	
3	elongacja	7 min	72°C	-

* zależna od temperatury topnienia użytych starterów - T_M obliczono wg wzoru
 $T_M = 2(A+T) + 4(G+C) - 5$ [°C]

Produkty reakcji PCR analizowano metodą elektroforezy w 1% żelu agarozowym (podrozdział 3.2.1.3.).

3.2.1.2. Oczyszczanie DNA

Produkty reakcji PCR oraz dsDNA po innych reakcjach enzymatycznych oczyszczano wykorzystując gotowy zestaw odczynników QIAquick PCR Purification Kit firmy QIAGEN, zgodnie z instrukcją producenta.

3.2.1.3. Rozdział elektroforetyczny kwasów nukleinowych w żelu agarozowym w warunkach natywnych

Do badanych preparatów dodawano buforu obciążającego 6x BJ do uzyskania jednokrotnego stężenia końcowego. Tak przygotowane próby nakładano na 1% żel agarozowy do rozdziału elektroforetycznego kwasów nukleinowych w warunkach

natywnych (żel przygotowano zgodnie z instrukcją zawartą w podrozdziale 3.1.8.). Rozdział prowadzono w buforze 0,5x TBE, przy napięciu 125V i w obecności wzorca mas (DirectLoad™ Wide Range DNA Marker, SIGMA).

3.2.1.4. Rozdział elektroforetyczny kwasów nukleinowych w żelu poliakrylamidowym w warunkach denaturujących

Do badanych preparatów dodawano buforu obciążającego 2x BB do uzyskania jednokrotnego stężenia końcowego. Próby denaturowano przez inkubację 5 min w temperaturze 95°C i gwałtowne schłodzenie na lodzie. Tak przygotowane próby nakładano na 15% żel poliakrylamidowy do rozdziału elektroforetycznego kwasów nukleinowych w warunkach denaturujących (żel przygotowano zgodnie z instrukcją zawartą w podrozdziale 3.1.8.). Rozdział prowadzono w buforze 1x TBE, przy natężeniu prądu 10 mA przez pierwsze 30 min elektroforezy (tzw. preelektroforeza), a następnie przy natężeniu prądu 40 mA.

3.2.1.5. Wizualizacja i oznaczanie stężenia kwasów nukleinowych

Kwasy nukleinowe rozdzielane w żelach agarozowych w warunkach natywnych wizualizowano w świetle UV wykorzystując zjawisko interkalacji bromku etydyny. Dokumentację żeli prowadzono z wykorzystaniem zestawu Gel Logic 200 firmy Kodak.

Znakowane radioizotopowo kwasy nukleinowe rozdzielane w żelach poliakrylamidowych w warunkach denaturujących wizualizowano za pomocą autoradiografii. Ekspozycję żeli prowadzono w temperaturze pokojowej stosując ekrany firmy Kodak, skanowane następnie przy pomocy skanera materiałów znakowanych radioizotopowo FLA-5100 Phosphorimager firmy Fujifilm. Ilościową i jakościową analizę rozdzielonych kwasów nukleinowych przeprowadzano przy pomocy programu Multi Gauge V3.0.

Stężenie DNA w próbkach oznaczano spektrofotometrycznie na podstawie pomiarów absorpcji światła UV o długości fali 260 nm w spektrofotometrze Nanodrop 2000 firmy Thermo Scientific.

3.2.1.6. Sekwencjonowanie DNA

Sekwencjonowanie DNA prowadzono z wykorzystaniem gotowego zestawu odczynników BigDye[®] Terminator V1.1 cycle sequencing kit firmy Applied Biosystems. Skład mieszaniny reakcyjnej przedstawiono w Tabeli 3.2.5. Warunki reakcji przedstawiono w Tabeli 3.2.6. Reakcję sekwencjonowania prowadzono w termocyklerze firmy Biometra.

Tabela 3.2.5. Skład mieszaniny reakcyjnej sekwencjonowania DNA.

odczynnik (stężenie początkowe)	ilość (stężenie końcowe)
matryca DNA	150 ng
starter	15 pmol
50% (v/v) DMSO	2 µl
<i>Ready Reaction Premix*</i>	1 µl
<i>5x BigDye Sequencing Buffer*</i>	3 µl
H ₂ O	do 20 µl

* dostarczone przez producenta zestawu odczynników BigDye[®] Terminator V1.1 cycle sequencing kit

Tabela 3.2.6. Warunki reakcji sekwencjonowania.

numer etapu	etap	czas trwania	temperatura	liczba powtórzeń
1	denaturacja	1 min	96°C	1
2	denaturacja	10 s	96°C	35
	hybrydyzacja	10 s	50°C	
	elongacja	4 min	60°C	

Oczyszczanie produktów reakcji sekwencjonowania

- Do mieszaniny reakcyjnej dodawano 5 µl 125 mM EDTA i 60 µl 96% etanolu.
- Wytrząsano i inkubowano 15 min w temperaturze pokojowej.
- Wirowano 30 min z prędkością 14000 rpm w temperaturze 4°C.
- Supernatant dekantowano, a osad przemywano 200 µl 70% etanolu.
- Wirowano 10 min z prędkością 14000 rpm w temperaturze 4°C.
- Supernatant dekantowano, a osad suszono.

- Osad rozpuszczano w 10 μ l formamidu.
- Tak przygotowane próbki dostarczano do Wydziałowej Pracowni Techniki Biologii Molekularnej Wydziału Biologii Uniwersytetu im. Adama Mickiewicza, gdzie dokonywano rozdzielania elektroforetycznego produktów reakcji i odczytu sekwencji.

3.2.1.7. Trawienie enzymami restrykcyjnymi produktów PCR oraz plazmidów

Cięcie restrykcyjne stosowano w celu przygotowania wektora i insertu do ligacji (podrozdział 3.2.1.8.) bądź też w celu analizy plazmidów izolowanych w małej skali (podrozdział 3.2.1.10.). Skład mieszaniny reakcyjnej standardowego cięcia enzymami restrykcyjnymi przedstawiono w Tabeli 3.2.7. Skład mieszaniny reakcyjnej cięcia enzymami restrykcyjnymi typu FastDigest[®] przedstawiono w Tabeli 3.2.8. Mieszaninę reakcyjną inkubowano 4 h lub 1 h (w przypadku enzymów typu FastDigest[®]) w temperaturze 37°C. Produkty cięcia analizowano w 1% żelu agarozowym (podrozdział 3.2.1.3.).

Tabela 3.2.7. Skład mieszaniny reakcyjnej standardowego cięcia enzymami restrykcyjnymi.

odczynnik	Ilość	
	skala preparatywna	skala analityczna
dsDNA	1 – 10 μ g	0,1 – 1 μ g
bufor* 10x	5 μ l	1 μ l
enzym	25 – 50 U	5 – 10 U
H ₂ O	do 50 μ l	do 10 μ l

* dostarczony przez producenta wraz z enzymem

Tabela 3.2.8. Skład mieszaniny reakcyjnej cięcia enzymami restrykcyjnymi typu FastDigest[®].

odczynnik	Ilość	
	skala preparatywna	skala analityczna
dsDNA	1 – 10 μ g	0,1 – 1 μ g
bufor* 10x <i>FastDigest</i>	2 μ l	1 μ l
enzym <i>FastDigest</i>	2 μ l	1 μ l
H ₂ O	do 20 μ l	do 10 μ l

* dostarczony przez producenta wraz z enzymem

3.2.1.8. Ligacja

Reakcję ligacji prowadzono z wykorzystaniem odczynników firmy Fermentas. Skład mieszniny reakcyjnej podano w Tabeli 3.2.9. Mieszaninę reakcyjną inkubowano 16 h w temperaturze 16°C.

Tabela 3.2.9. Skład mieszaniny reakcyjnej ligacji.

odczynnik	Ilość
wektor	50 – 100 ng
insert	do stosunku molowego insert/wektor równego 3/1
<i>10x Ligation Buffer</i> *	1 µl
ligaza DNA T4 (5 U/µl)	1 µl
H ₂ O	do 10 µl

* dostarczony przez producenta wraz z enzymem

3.2.1.9. Ligacja z wektorem pGEM-T Easy

Reakcję ligacji z wektorem pGEM-T Easy prowadzono z wykorzystaniem gotowego zestawu odczynników pGEM[®]-T Easy Vector System firmy Promega. Skład mieszaniny reakcyjnej podano w Tabeli 3.2.10. Mieszaninę reakcyjną inkubowano 16 h w temperaturze 4°C.

Tabela 3.2.10. Skład mieszaniny reakcyjnej ligacji z wektorem pGEM-T Easy.

odczynnik	Ilość
wektor	50 ng
insert	do stosunku molowego insert/wektor równego 3/1
<i>2x Rapid Ligation Buffer</i> *	5 µl
ligaza DNA T4 (3 U/µl)	1 µl
H ₂ O	do 10 µl

* dostarczony przez producenta wraz z enzymem

3.2.1.10. Izolacja plazmidów w małej skali (do 10 µg)

- Pojedyncze kolonie bakterii (zawierające pożądany plazmid) wyhodowane na stałej pożywce LB przenoszono do 2 ml płynnej pożywki LB zawierającej odpowiedni antybiotyk¹³ i inkubowano 16 h w temperaturze 37°C, wytrząsając z prędkością 250 rpm.

¹³ Antybiotyki dobrane były odpowiednio do genów oporności danego szczepu bakterii (Tabela 3.1.6.) i genu oporności zawartym w plazmidzie, którym dany szczep bakterii był transformowany (podrozdział 3.1.3.).

- 1,5 ml hodowli wirowano 1 min z prędkością 14000 rpm w temperaturze pokojowej.
- Supernatant dekantowano, a zawierający bakterie osad zawieszano w 100 μ l roztworu ALM1, wytrząsając 5 min w temperaturze pokojowej.
- Dodawano 200 μ l świeżo przygotowanego roztworu ALM2, delikatnie mieszano i inkubowano 5 min na lodzie.
- Dodawano 150 μ l 7,5 M roztworu octanu amonu, mieszano i inkubowano 20 min na lodzie.
- Wirowano 20 min z prędkością 14000 rpm w temperaturze pokojowej.
- Supernatant przenoszono do nowej probówki, dodawano 1 ml 96% etanolu i inkubowano 20 min na lodzie.
- Wirowano 15 min z prędkością 14000 rpm w temperaturze 4°C.
- Supernatant dekantowano, a osad suszono.
- Osad rozpuszczano w roztworze zawierającym: 49,5 μ l H₂O i 0,5 μ l RNazyA o stężeniu 10 μ g/ μ l.
- Otrzymane plazmidy poddawano analizie restrykcyjnej (podrozdział 3.2.1.7.) lub analizie z wykorzystaniem reakcji PCR (podrozdział 3.2.1.1.). Wybrane plazmidy oczyszczano na kolumnkach QIAquick firmy QIAGEN, zgodnie z instrukcją producenta.

3.2.1.11. Izolacja plazmidów w średniej skali (do 100 μ g)

Izolację plazmidów w średniej skali prowadzono z wykorzystaniem gotowego zestawu odczynników: Plasmid Midi Kit firmy Qiagen zgodnie z instrukcją producenta.

3.2.2. Metody rozdziału i analizy białek

3.2.2.1. Rozdział elektroforetyczny białek w żelu poliakrylamidowym w warunkach denaturujących (żel PAA)

- Przygotowanie żelu

Żel PAA do rozdziału elektroforetycznego białek składał się z 5% żelu zagęszczającego i 12 lub 15% żelu rozdzielającego (zele przygotowywano zgodnie z instrukcją zawartą w podrozdziale 3.1.8). Żel wylewano między szyby i umieszczano w aparacie Mini-Protean firmy BioRad wypełnionym jednokrotnie stężonym buforem R.

- Przygotowanie i rozdział prób

Do badanych preparatów dodawano 4x stężonego buforu obciążającego SB do uzyskania jednokrotnego stężenia końcowego. Próby denaturowano przez inkubację 5 min w temperaturze 95°C. Tak przygotowane próby rozdzielano w żelu PAA w obecności wzorca mas (Novex[®] Sharp Pre-Stained Protein Standards, INVITROGEN). Rozdział prowadzono przy napięciu 80 V dla żelu zagęszczającego i 130 V dla żelu rozdzielającego.

- Wizualizacja produktów rozdziału

Produkty rozdziału uwidaczniano przez podgrzanie żelu w kuchence mikrofalowej w roztworze barwiącym, a następnie kilkugodzinne płukanie żelu w temperaturze pokojowej w roztworze odbarwiającym. Żele dokumentowano używając zestawu Gel Logic 200 firmy Kodak.

3.2.2.2. Identyfikacja białek metodą *western blot*

- Rozdział elektroforetyczny badanego preparatu białkowego w żelu PAA w warunkach denaturujących

Badany preparat białkowy rozdzielano w żelu PAA (zgodnie z instrukcją zawartą w podrozdziale 3.2.2.1.). Żelu po rozdziale nie poddawano barwieniu.

- Elektrotransfer rozdzielonych białek na membranę PVDF

Z żelu po rozdziale odcinano część zagęszczającą, pozostawiając jedynie część rozdzielającą. Membranę PVDF (przyciętą do rozmiaru żelu) zanurzano w metanolu na 2 min. Żel i 6 kawałków bibuły typu Whatman (przyciętych do rozmiaru żelu) moczone w buforze do elektrotransferu. W aparacie do elektrotransferu metodą półsuchą firmy BioRad układano kolejno: 3 bibuły typu Whatman, membranę PVDF, żel i pozostałe 3 bibuły. Elektrotransfer prowadzono przez 1 h przy natężeniu prądu równemu co do wartości polu powierzchni membrany oraz napięciu nieprzekraczającemu 10V.

- Detekcja białka na membranie

Membranę umieszczano w urządzeniu SNAP i.d.TM firmy Millipore (połączonym z pompą próżniową) lokując ją w przeznaczony do tego celu kasecie (zgodnie z instrukcją producenta). Blokowanie membrany prowadzono przepuszczając przez kasetę z membraną (pod ciśnieniem) 15 ml roztworu blokującego. Następnie membranę inkubowano 10 min z 2,5 ml roztworu przeciwciała I-rzędowego, po czym płukano trzykrotnie przepuszczając 20 ml buforu TBSTT przez kasetę z membraną. Membranę inkubowano 10 min z 2,5 ml roztworu przeciwciała II-rzędowego i trzykrotnie płukano

20 ml buforu TBSTT. Membranę wyciągano z kasety i urządzenia SNAP i.d.TM. Inkubowano kilka minut w roztworze do detekcji kolorymetrycznej. Natychmiast po pojawieniu się barwnych prążków, membranę przekładano do nowego pojemnika, płukano H₂O i pozostawiono do wyschnięcia. Membrany dokumentowano używając skanera komputerowego.

3.2.3. Otrzymywanie AID i jej mutantów w systemie prokariotycznym

3.2.3.1. Klonowanie genu AID do wektora ekspresyjnego

DNA kodujący ludzkie białko AID poddawano ligacji (podrozdział 3.2.1.8.) z wektorem ekspresyjnym pGEX-4T-1 lub pGEX-6p-3 otrzymując w ten sposób konstrukt ekspresyjny. Plazmid pCDFDuet-1 z wklonowanym DNA kodującym ludzką AID otrzymano dzięki uprzejmości dr Niny Papavasiliou z The Rockefeller University (Nowy Jork, USA). Wymienione konstrukty ekspresyjne służyły do produkcji białka AID w komórkach bakteryjnych.

3.2.3.2. Otrzymywanie zmutowanych wariantów AID

DNA kodujący wariant AID skrócony o 10 aminokwasów z końca C – mutant Δ189 otrzymywano metodą PCR. Jako matrycę zastosowano plazmid zawierający DNA kodujący białko AID typu dzikiego. W reakcji użyto startery: F_04_aid oraz mutΔ189_R (Tabela 3.5.). Skład mieszaniny reakcyjnej przedstawiono w Tabeli 3.2.3. Warunki reakcji PCR opisano w Tabeli 3.2.4. Powielony fragment genu AID poddawano cięciu restrykcyjnemu i ligacji z wektorem ekspresyjnym (analogicznie jak w przypadku DNA kodującego AID typu dzikiego).

Wektory kodujące warianty AID zawierające mutacje punktowe: R50A i N51A otrzymywano na drodze ukierunkowanej mutagenyzy. Metoda ta polega na amplifikacji całego plazmidu (bez potrzeby wycinania i ligowania modyfikowanych odcinków) przy użyciu dwóch wzajemnie komplementarnych starterów zawierających odpowiednie mutacje – startery mutR50A_F i mutR50A_R dla mutantu R50A oraz startery mutN51A_F i mutN51A_R dla mutantu N51A (Tabela 3.5.). Jako matrycę dla tej reakcji wykorzystywano wektor kodujący AID typu dzikiego. Do amplifikacji plazmidu wykorzystywano polimerazę o wysokiej wierności kopiowania – polimerazę DNA Pfu. Skład mieszaniny reakcyjnej oraz warunki reakcji przedstawiono w Tabeli 3.2.11. i 3.2.12.

Tabela 3.2.11. Skład mieszaniny reakcyjnej stosowanej w ukierunkowanej mutagenizie.

odczynnik (stężenie początkowe)	ilość (stężenie końcowe)
matryca DNA (wektor kodujący AID typu dzikiego)	300 ng
jeden ze starterów: starter F lub starter R	40 pmol
bufor reakcyjny 10x (dostarczony wraz z polimerazą Pfu przez firmę Promega)	5 μ l
mix dNTPs (10 mM każdy)	1 μ l (0,2 mM)
polimeraza DNA Pfu (3 U/ μ l)	0,5 μ l (1,5 U)
H ₂ O	do 50 μ l

- Przygotowywano dwie mieszaniny reakcyjne (Tabela 3.2.11.). Każda z nich zawierała jeden z pary starterów.
- Prowadzono reakcję (Tabela 3.2.12.) przez 13 cykli w termocyklerze firmy Biometra.
- Produkty reakcji łączono, mieszano a następnie rozdzielano do dwóch nowych probówek. Do każdej dodawano 3 U polimerazy DNA Pfu.
- Prowadzono reakcję (Tabela 3.2.12.) przez kolejne 25 cykli.
- Produkty reakcji oczyszczano (podrozdział 3.2.1.2.), a następnie poddawano trawieniu enzymem DpnI inkubując 4 h w temperaturze 37°C następującą mieszaniną reakcyjną:
 - ✓ 5 - 10 μ g oczyszczonego DNA
 - ✓ 10 μ l buforu 10x (dostarczonego wraz z enzymem DpnI przez producenta)
 - ✓ 60 U enzymu DpnI
 - ✓ dopełniano H₂O do 100 μ l
- Produkty reakcji trawienia ponownie oczyszczano, a uzyskanym DNA transformowano komórki kompetentne *E.coli* szczepu One shot[®] TOP10 (podrozdział 3.2.3.3.). Z uzyskanych kolonii bakteryjnych izolowano plazmidy w małej skali (podrozdział 3.2.1.10.), poddawano je analizie restrykcyjnej (podrozdział 3.2.1.7.) i sekwencjonowano (podrozdział 3.2.1.6.). Zidentyfikowano plazmidy zawierające oczekiwaną mutację.

Jeden z nich posłużył do transformacji szczepu bakterii stosowanego do produkcji białka (podrozdział 3.2.3.3.).

Tabela 3.2.12. Warunki reakcji ukierunkowanej mutagenazy.

numer etapu	Etap	czas trwania	temperatura	liczba cykli
1	Denaturacja	2 min	96°C	-
2	Denaturacja	50 s	96°C	13 cykli, następnie 25 cykli
	hybrydyzacja	50 s	60°C	
	elongacja	12 min	68°C	
3	elongacja	12 min	68°C	-

3.2.3.3. Transformacja komórek bakteryjnych metodą szoku cieplnego

- Do zawiesiny komórek kompetentnych *E. coli* dodawano 10 ng oczyszczonego plazmidu lub 5 µl produktu ligacji.
- Inkubowano 5 min na lodzie.
- Zawiesinę ogrzewano 45 s (szczep DH5α) lub 30 s (pozostałe szczepy) w temperaturze 42°C.
- Inkubowano 5 min na lodzie.
- Do transformowanych komórek dodawano 1 ml płynnej pożywki LB (szczep DH5α) lub 250 µl pożywki SOC (pozostałe szczepy) i inkubowano 1 h w temperaturze 37°C.
- W przypadku transformacji szczepu DH5α lub szczepu One shot® TOP10 plazmidem pGEM®-T Easy (umożliwiającym selekcję komórek zawierających wektor z insertem na podstawie testu α-komplementacji) do 0,1 – 0,3 ml otrzymanej zawiesiny dodawano 15 µl 0,1 M IPTG oraz 30 µl X-Gal.
- Otrzymaną zawiesinę bakterii rozprowadzano na szalkach ze stałą pożywką LB zawierającą antybiotyk zgodny z genem oporności kodowanym przez plazmid, którym transformowane były bakterie (podrozdział 3.1.3.).
- Szalki inkubowano 16 h w temperaturze 37°C.

3.2.3.4. Wytwarzanie białka AID w komórkach bakteryjnych

Każdy z przygotowanych konstruktów ekspresyjnych (plazmid zawierający DNA kodujący białko AID) wprowadzano metodą szoku termicznego (podrozdział 3.2.3.3.) do odpowiedniego szczepu bakterii do produkcji białka. Zestawienie

poszczególnych układów ekspresyjnych stosowanych do produkcji białka AID przedstawiono w Tabeli 3.2.13.

Tabela 3.2.13. Zastosowane układy ekspresyjne oraz odpowiadające im składy pożywek do hodowli bakterii.

szczep <i>E. coli</i>	plazmid	Pożywka
BL21(DE3)pLysS	pGEX-4T-1	LB + chloramfenikol* + ampicylina
	pGEX-6p-3	LB + chloramfenikol + ampicylina
	pCDFDuet-1	LB + chloramfenikol + spektynomycyna
Origami(DE3)pLysS	pGEX-4T-1	LB + chloramfenikol + kanamycyna + ampicylina
BL21-AL TM OneShot [®]	pGEX-4T-1	LB + ampicylina
Rosetta(DE3)pLysS	pGEX-4T-1	LB + chloramfenikol + ampicylina

* stosowano następujące stężenia końcowe antybiotyków: chloramfenikol 34 µg/ml, ampicylina 200 µg/ml, spektynomycyna 50 µg/ml, kanamycyna 30 µg/ml.

3.2.3.5. Produkcja AID i jej mutantów w komórkach bakteryjnych w skali analitycznej i przygotowanie stoków glicerolowych

Doświadczenia w skali analitycznej prowadzono w celu ustalenia optymalnych warunków hodowli oraz przygotowania preparatów do analizy rozpuszczalności białek.

- Z pojedynczej kolonii bakterii wyhodowanych na szalce ze stałą pożywką LB (po transformacji plazmidem ekspresyjnym kodującym AID, podrozdział 3.2.3.3.) pobierano zaszczepkę i przenoszono do 2 ml płynnej pożywki LB zawierającej odpowiedni antybiotyk (Tabela 3.2.13). Hodowlę inkubowano 16 h w temperaturze 37°C wytrząsając z prędkością 250 rpm.
- Pobierano 200 µl hodowli i zaszczepiano nią 20 ml świeżej pożywki LB zawierającej odpowiedni antybiotyk (Tabela 3.2.13). Hodowlę inkubowano w temperaturze 37°C wytrząsając z prędkością 250 rpm do momentu osiągnięcia gęstości optycznej $OD_{600} = 0,6 - 0,8$.
- Pobierano 1 ml kultury bakteryjnej. Próbę wirowano 1 min z prędkością 12000 rpm w temperaturze pokojowej, roztwór dekantowano, osad zawieszano w 100 µl buforu 1x SB i zamrażano w temperaturze -20°C. Po rozmrożeniu próbę wirowano 30 min z prędkością 14000 rpm w temperaturze 4°C.
- Z niewielkiej ilości hodowli sporządzano stoki glicerolowe: pobierano 800 µl kultury bakteryjnej, dodawano 100 µl 80% glicerolu, mieszano

i zamrażano w temperaturze -80°C . Tak przygotowane stoki glicerolowe wykorzystywano podczas produkcji białka w skali preparatywnej.

- W pozostałej części hodowli bakteryjnej indukowano ekspresję genu kodującego białko AID przez dodanie do pożywki 1 M IPTG do uzyskania stężenia końcowego 0,5 mM. W przypadku szczepu bakterii BL21-ALTMOneShot[®] ekspresję genu kodującego białko AID indukowano poprzez dodanie do pożywki 1 M IPTG do uzyskania stężenia końcowego 1 mM oraz 20% L-arabinozy do uzyskania stężenia końcowego 0,2%.
- Inkubowano hodowlę 3 h w temperaturze 37°C lub 16 h w temperaturze 18°C wytrząsając z prędkością 250 rpm. W przypadku 3 h hodowli, co godzinę pobierano 1 ml kultury bakteryjnej. W przypadku 16 h hodowli pobierano 1 ml kultury bakteryjnej po 16 h.
- Próby wirowano 1 min z prędkością 12000 rpm w temperaturze pokojowej, roztwór dekantowano, osady zawieszano w 100 μl buforu 1x SB i zamrażano w temperaturze -20°C . Po rozmrożeniu próby wirowano 30 min z prędkością 14000 rpm w temperaturze 4°C .
- Otrzymane po wirowaniu supernatany (prób pobranych przed i po indukcji ekspresji genu AID) analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.).
- Po upływie 3 lub 16 h pobierano dodatkowo dwie próby po 1 ml zawiesiny w celu analizy rozpuszczalności białek (podrozdział 3.2.3.6.).

3.2.3.6. Analiza rozpuszczalności białek

- Pobierano 1 ml kultury bakteryjnej po upływie 3 h lub 16 h od indukcji ekspresji genu AID (hodowla bakterii prowadzona w celu produkcji AID w skali analitycznej, podrozdział 3.2.3.5.). Próbkę wirowano 1 min z prędkością 12000 rpm w temperaturze pokojowej, roztwór dekantowano, a osad zamrażano w temperaturze -20°C .
- Po rozmrożeniu osad zawieszano w 100 μl buforu A. W przypadku szczepu bakterii BL21-ALTMOneShot[®] do buforu A dodawano lizozymu do uzyskania stężenia końcowego 2 mg/ml.
- Zawiesinę inkubowano 15 min na lodzie, a następnie wirowano 15 min z prędkością 12000 rpm w temperaturze 4°C .

- Supernatant, zawierający frakcję białek rozpuszczalnych, przenoszono do nowej probówki, dodawano buforu 4x SB do uzyskania jednokrotnego stężenia końcowego i analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.).
- Osad zawierający frakcję białek nierozpuszczalnych zawieszano w 100 μ l buforu B i inkubowano 15 min w temperaturze 37°C.
- Wirowano 15 min z prędkością 12000 rpm w temperaturze 4°C.
- Supernatant, zawierający frakcję białek nierozpuszczalnych, przenoszono do nowej probówki, dodawano buforu 4x SB do uzyskania jednokrotnego stężenia końcowego i analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.).

3.2.3.7. Produkcja AID i jej mutantów w komórkach bakteryjnych w skali preparatywnej

- Ze stoku glicerolowego bakterii transformowanych plazmidem ekspresyjnym kodującym AID (lub też AID z wprowadzoną mutacją) pobierano 300 μ l i zaszczepiano 20 ml pożywki LB zawierającej odpowiedni antybiotyk (Tabela 3.2.13).
- Zawiesinę bakterii inkubowano 16 h w temperaturze 37°C wytrząsając z prędkością 250 rpm.
- Uzyskaną kulturą zaszczepiano 1 l świeżej pożywki LB zawierającej odpowiedni antybiotyk (Tabela 3.2.13). Hodowlę wytrząsano z prędkością 250 rpm w temperaturze 37°C do momentu osiągnięcia gęstości optycznej $OD_{600} = 0,6 - 0,8$.
- Indukowano ekspresję genu kodującego białko AID przez dodanie do pożywki 1 M IPTG do uzyskania stężenia końcowego 0,5 mM. W przypadku szczepu bakterii BL21-ALTMOneShot[®] ekspresję genu kodującego białko AID indukowano poprzez dodanie do pożywki 1 M IPTG do uzyskania stężenia końcowego 1 mM oraz 20% L-arabinozy do uzyskania stężenia końcowego 0,2%.
- Hodowlę bakteryjną inkubowano 3 h w temperaturze 37°C lub 16 h w temperaturze 18°C wytrząsając z prędkością 250 rpm.
- Zawiesinę bakteryjną wirowano 15 min z prędkością 4000 rpm w temperaturze 4°C.

- Supernatant dekantowano, a osad zamrażano w temperaturze -20°C .
- Zamrożony osad bakteryjny używano do izolacji i oczyszczania zawartego w nim białka AID (podrozdział 3.2.3.8.).

3.2.3.8. Oczyszczanie białek otrzymanych w komórkach bakteryjnych

Oczyszczanie białek na kolumnie wypełnionej złożem glutationowym

- Osad komórek bakteryjnych (podrozdział 3.2.3.7.), z 1 l hodowli bakteryjnej prowadzonej w skali preparatywnej, zawieszano w 50 ml buforu 1x PBS zawierającego inhibitor proteaz, DTT w stężeniu końcowym 1 mM, oraz Triton X-100 w stężeniu końcowym 1%.
- Zawiesinę inkubowano 30 min na lodzie, sonikowano trzykrotnie przez 15 s (procent amplitudy drgań fal dźwiękowych = 70), a następnie wirowano 15 min z prędkością 12000 rpm w temperaturze 4°C .
- Do zawartego w supernatancie ekstraktu białek rozpuszczalnych dodawano 2 ml złoża glutationowego zrównoważonego buforem 1x PBS.
- Inkubowano 30 min w temperaturze pokojowej delikatnie wytrząsając.
- Zawiesinę nakładano na kolumnę chromatograficzną i zbierano frakcję białek niezwiązanych ze złożem.
- Przemywano złożę 30 ml buforu 1x PBS zawierającego DTT w stężeniu końcowym 1 mM, w celu odmycia reszty białek niezwiązanych.
- Eluowano frakcję białek związanych ze złożem przemywając kolumnę pięciokrotnie 2 ml buforu do elucji.
- Zebrane frakcje analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.).

Oczyszczanie białek na kolumnie wypełnionej złożem heparynowym

- Preparat białkowy, powstały poprzez połączenie frakcji eluowanych z kolumny wypełnionej złożem glutationowym, łączono z 2 ml złoża heparynowego zrównoważonego 50 mM Tris-HCl, pH = 8.
- Inkubowano 30 min w temperaturze pokojowej delikatnie wytrząsając.
- Zawiesinę nakładano na kolumnę chromatograficzną i zbierano frakcję białek niezwiązanych ze złożem.
- Kolumnę przemywano gradientem NaCl w 50 mM Tris-HCl, pH = 8. Na kolumnę nakładano 5 ml porcji buforu zawierającego NaCl

w następujących stężeniach: 50, 130, 250, 600, 2000 mM. Zbierano frakcje białek związanych ze złożem.

- Zebrane frakcje zagęszczano do objętości 300 μ l (podrozdział 3.2.3.9.) i analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.).

Oczyszczanie białek na kolumnie wypełnionej złożem niklowym

Chromatografię powinowactwa do jonów Ni^{2+} stosowano jako pierwszą bądź też drugą w kolejności metodę oczyszczania białek. W tym drugim przypadku pomijano pierwsze dwa etapy, a ekstrakt białkowy przygotowywano poprzez wymianę w nim buforu na bufor I (podrozdział 3.2.3.9.).

- Osad komórek bakteryjnych (podrozdział 3.2.3.7.) z 1 l hodowli bakteryjnej prowadzonej w skali preparatywnej zawieszano w buforze I zawierającym inhibitor proteaz. Na 1 g osadu bakteryjnego stosowano 2 ml buforu.
- Zawiesinę inkubowano 30 min na lodzie, sonikowano trzykrotnie przez 15 s (procent amplitudy drgań fal dźwiękowych = 70), dodawano 0,2 ml 4 M NaCl na 1 g osadu, a następnie wirowano 15 min z prędkością 12000 rpm w temperaturze 4°C.
- Do supernatantu zawierającego ekstrakt białek rozpuszczalnych dodawano jedną objętość buforu II i 1 ml złoża niklowego, zrównoważonego buforem II.
- Inkubowano 30 min w temperaturze pokojowej delikatnie wytrząsając.
- Zawiesinę nakładano na kolumnę chromatograficzną zbierając frakcje białek niezwiązanych ze złożem.
- Kolumnę przemywano 10 ml buforu II.
- Kolumnę przemywano gradientem imidazolu w jednokrotnie stężonym buforze III. Na kolumnę nakładano 10 ml porcje buforu 1x III zawierającego imidazol w następujących stężeniach: 10, 20, 30, 40, 50, 60, 100, 200 mM. Zbierano frakcje o objętości 2,5 ml zawierające białka związane ze złożem. Przemywano kolumnę 10 ml 1 M imidazolu.
- Zebrane frakcje analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.).

3.2.3.9. Wymiana buforu i zagęszczanie preparatów białkowych

Preparaty białkowe zagęszczano przy użyciu filtrów Amicon Ultra firmy Millipore (o objętości 4 lub 15 ml i punkcie odcięcia 10 000 MWC) zgodnie z instrukcją producenta. Podczas wymiany buforu, preparat białkowy cyklicznie zagęszczano, a następnie uzupełniano objętość nowym buforem. Czynności powtarzano do momentu przemycia preparatu 30 objętościami nowego buforu.

3.2.3.10. Sączenie molekularne

- Preparat białkowy (wstępnie oczyszczony na kolumnie wypełnionej złożem glutationowym) poddawano wymianie buforu na bufor do sączenia molekularnego SE1 lub SE2 i zagęszczaniu do objętości 4 ml.
- Tak przygotowany preparat nakładano na kolumnę wypełnioną 45 ml złoża Superdex75 16/60 HiLoad firmy GE Healthcare zrównoważonego buforem SE1 lub SE2.
- Kolumnę przemywano buforem SE1 lub SE2 z prędkością przepływu wynoszącą 0,375 ml/min, zbierając frakcje o objętości 2,5 ml do momentu wypłukania wszystkich białek z kolumny.
- Zebrane frakcje analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.).

3.2.3.11. Oznaczanie stężenia białka metodą Bradford

- Do 5 μ l badanego preparatu białkowego (o nieznanym stężeniu) dodawano 250 μ l odczynnika Bradford (Fermentas) i inkubowano 5 min w temperaturze pokojowej. Dla każdej próby wykonywano trzy powtórzenia techniczne.
- Analogiczne mieszaniny sporządzano dla roztworów wzorcowych białka BSA o stężeniu końcowym: 0,25; 0,5; 1; 2 mg/ml oraz próby ślepej, niezawierającej białka (zawierającej jedynie bufor, w którym znajdowało się badane białko).
- Dokonywano pomiaru absorpcji światła o długości fali 595 nm z wykorzystaniem spektrofotometru BioPhotometer firmy Eppendorf.
- Na podstawie uzyskanych wyników dla prób wzorcowych wykreślano krzywą wzorcową. Stężenie badanego preparatu białkowego odczytywano

z krzywej wzorcowej na podstawie średniej arytmetycznej z wartości absorpcji dla próby badanej.

3.2.4. Metody analizy aktywności enzymatycznej AID

3.2.4.1. Radioizotopowe znakowanie DNA na końcu 5'

Oligodeoksyrybonukleotydy stosowane w badaniach aktywności AID (Tabela 3.5.) znakowano radioizotopowo na końcu 5' za pomocą $\gamma^{[32]P}$ ATP w reakcji katalizowanej przez kinazę polinukleotydową T4 firmy Fermentas, zgodnie z instrukcją producenta. Skład mieszaniny reakcyjnej oraz warunki reakcji przedstawiono w Tabeli 3.2.14.

Tabela 3.2.14. Skład mieszaniny reakcyjnej oraz warunki reakcji znakowania DNA na końcu 5'.

odczynnik	ilość
DNA	20 pmol
10x <i>PNK buffer A</i> *	2 μ l
H ₂ O	do 18 μ l
denaturowano przez inkubację 5 min w temperaturze 95°C i nagle schłodzenie na lodzie. Następnie dodawano:	
kinazę polinukleotydową T4 (10 U/ μ l)	1 μ l
$\gamma^{[32]P}$ dATP (3000 Ci/mmol)	1 μ l
mieszaninę inkubowano 20 min w temperaturze 37°C	
mieszaninę inkubowano 10 min w temperaturze 75°C	

* dostarczony przez firmę Fermentas wraz z kinazą polinukleotydową T4

- Produkty reakcji znakowania oczyszczano wykorzystując kolumnienki NucAwayTM Spin Columns firmy Ambion, zgodnie z instrukcją producenta.
- Pomiaru radioaktywności dokonywano z wykorzystaniem licznika scyntylacyjnego i czytnika mikroplętek MicroBeta2TM firmy PerkinElmer.

3.2.4.2. Badanie aktywności deaminazowej AID wobec pojedynczego motywu WRCY

W celu wykazania, że produkowane białko wykazuje aktywność deaminazową wobec pojedynczego motywu WRCY stosowano metodę opublikowaną w 2003 roku przez Sohail i współautorów [176]. Jako substratu do reakcji używano 80-nukleotydowej cząsteczki ssDNA (Oligo_C, Tabela 3.5.) znakowanej radioizotopowo na końcu 5' (podrozdział 3.2.4.1.). Cząsteczka ta zawiera tylko jedną resztę cytydynową w pozycji 40 oligonukleotydu.

- 0,3 pmola znakowanej radioizotopowo cząsteczki ssDNA inkubowano 30 minut w temperaturze 37°C z testowanym preparatem białkowym (w ilości 4,5 µg całkowitego białka) w jednokrotnie stężonym buforze reakcyjnym dla białka AID. Reakcję prowadzono w objętości 10 µl. W przypadku obecności aktywnej deaminazy cytydyny na tym etapie dochodzi do deaminacji cytydyny do urydyny.
- Do reakcji dodawano 2 U glikozylazy UDG (Fermentas) i inkubowano dalsze 30 minut w temperaturze 37°C. Enzym ten usuwa uracyl, pozostawiając miejsce apirymidynowe.
- Przeprowadzano lizę alkaliczną, poprzez dodawanie do reakcji NaOH do stężenia końcowego 0,2 M. Mieszaninę inkubowano 15 minut w temperaturze 80°C. W przypadku istnienia miejsca apirymidynowego dochodzi do alkalicznej hydrolizy wiązania fosfodiesterowego sąsiadującego z tym miejscem.
- Produkty reakcji analizowano metodą elektroforezy w 15% żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.1.4.). W przypadku obecności aktywności deaminazowej obserwuje się produkt o długości 40 nukleotydów.
- Równocześnie przeprowadzano dwie reakcje kontrolne: pozytywną reakcję kontrolną, bez preparatu białkowego, ale z wykorzystaniem cząsteczki substratowej posiadającej U zamiast C w pozycji 40 oligonukleotydu (Oligo_U, Tabela 3.5) oraz negatywną reakcją kontrolną, bez preparatu białkowego, ale z wykorzystaniem standardowej cząsteczki substratowej (Oligo_C).
- Podczas badania aktywności AID testowano cztery różne mieszaniny reakcyjne. Ich skład, jak również skład mieszanin reakcyjnych dla reakcji kontrolnych zawiera Tabela 3.2.15.

Tabela 3.2.15. Skład mieszanin reakcyjnych stosowanych podczas badania aktywności deaminazowej AID wobec pojedynczego motywu WRCY.

odczynnik	skład testowanych mieszanin reakcyjnych (1-4) i mieszanin reakcyjnych dla reakcji kontrolnych (K+, K-)					
	1.	2.	3.	4.	K+	K-
ZnCl ₂ (1mM)	—	—	80 μM	80 μM	80 μM	80 μM
RNaza A (10 μg/μl)	—	1 μg	—	1 μg	1 μg	1 μg
Oligo_C	0,3 pmol	0,3 pmol	0,3 pmol	0,3 pmol	—	0,3 pmol
Oligo_U	—	—	—	—	0,3 pmol	—
1x bufor reakcyjny dla AID	do 10 μl	do 10 μl	do 10 μl	do 10 μl	do 10 μl	do 10 μl
preparat białkowy zawierający AID	4,5 μg*	4,5 μg	4,5 μg	4,5 μg	4,5 μg	4,5 μg

* całkowitego białka

3.2.4.3. Badanie aktywności deaminazowej AID wobec pojedynczego motywu WR5mCY

W celu sprawdzenia czy produkowane białko wykazuje aktywność deaminazową wobec pojedynczego motywu WR5mCY stosowano zmodyfikowaną wersję metody badania aktywności AID wobec motywu WRCY [177]. Jako substratu do reakcji używano 80-nukleotydowej cząsteczki ssDNA (Oligo_5mC, Tabela 3.5.) znakowanej radioizotopowo na końcu 5' (podrozdział 3.2.4.1.) zawierającej tylko jedną resztę 5-metylocytydynową w pozycji 40 oligonukleotydu.

- 0,3 pmola znakowanej radioizotopowo cząsteczki ssDNA inkubowano 30 minut w temperaturze 37°C z testowanym preparatem białkowym (w ilości 4,5 μg całkowitego białka) w jednokrotnie stężonym buforze reakcyjnym dla białka AID. Reakcję prowadzono w objętości 10 μl. W przypadku obecności aktywności deaminazowej na tym etapie dochodzi do deaminacji 5-metylocytydyny do tymidyny.
- Do reakcji dodawano KCl do stężenia końcowego 50 mM oraz 30 pmola oligodeksyrybonukleotydu komplementarnego do badanej cząsteczki substratowej (Oligo_C-komplement, Tabela 3.5.). Mieszaninę reakcyjną inkubowano 5 min w temperaturze 95°C, a następnie powoli (1°C / min) schładzano do temperatury 60°C. W sytuacji gdy dochodzi do deaminacji uzyskiwany jest dsDNA z pojedynczym niesparowaniem G:T w pozycji 40 oligonukleotydu.

- Do reakcji dodawano 4,5 U glikozylazy TDG (Trevigene) i 2 µl 10x stężonego buforu (dostarczony przez producenta wraz z enzymem) do jednokrotnego stężenia końcowego. Reakcję prowadzono w objętości 20 µl. Inkubowano 2 h w temperaturze 60°C. W przypadku obecności pojedynczego niesparowania G:T dochodzi do depirymidynacji T nietworzącego komplementarnej pary z G (enzym TDG usuwa tyminę, pozostawiając miejsce apirymidynowe).
- Prowadzono lizę alkaliczną, poprzez dodanie do reakcji NaOH do stężenia końcowego 0,2 M. Mieszaninę inkubowano 15 minut w temperaturze 80°C. W przypadku istnienia miejsca apirymidynowego dochodzi do alkalicznej hydrolizy wiązania fosfodiesterowego w miejscu depirymidynacji.
- Produkty reakcji analizowano metodą elektroforezy w 15% żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.1.4.). W przypadku obecności aktywności deaminazowej obserwuje się produkt o długości 40 nukleotydów.
- Równocześnie przeprowadzano dwie reakcje kontrolne: pozytywną reakcję kontrolną, bez preparatu białkowego, ale z wykorzystaniem cząsteczki substratowej posiadającej T zamiast 5mC w pozycji 40 oligonukleotydu (Oligo_T, Tabela 3.5) oraz negatywną reakcją kontrolną, bez preparatu białkowego, ale z wykorzystaniem standardowej cząsteczki substratowej (Oligo_5mC).

3.2.4.4. Badanie aktywności deaminazowej AID wobec pojedynczego motywu **WR5hmCY**

W celu sprawdzenia czy produkowane białko wykazuje aktywność deaminazową wobec pojedynczego motywu **WR5hmCY** stosowano zbliżoną metodę do metody badania aktywności AID wobec motywu **WRCY**. Reakcja przebiegała analogicznie z uwzględnieniem poniższych różnic:

- Jako substratu do reakcji używano 80-nukleotydowej cząsteczki ssDNA (Oligo_5hmC, Tabela 3.5.) zawierającej tylko jedną resztę 5-hydroksymetylocytydynową w pozycji 40 oligonukleotydu.
- Glikozylazę UDG w drugim etapie reakcji zastąpiono glikozylazą hSMUG1 (New England Biolabs) (dodawano 5 U).

- W pozytywnej reakcji kontrolnej wykorzystywano cząsteczkę substratową posiadającą 5hmU zamiast 5hmC w pozycji 40 oligonukleotydu (Oligo_5hmU, Tabela 3.5).

Produkty reakcji analizowano metodą elektroforezy w 15% żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.1.4.). W przypadku obecności aktywności deaminazowej spodziewano się zaobserwować produkt o długości 40 nukleotydów.

3.2.4.5. Badanie aktywności deaminazowej AID na matrycach zawierających wszystkie sekwencje typu NNCN

- Zsyntetyzowano na drodze chemicznej 4 oligodeksynukleotydy zawierające, występujące kolejno po sobie wszystkie sekwencje typu NNCN, gdzie N oznacza dowolny nukleotyd (oligonukleotydy NNCN_1, NNCN_2, NNCN_3 lub NNCN_4, Tabela 3.5.).
- 0,3 pmola oligodeoksyrybonukleotydu zawierającego wiele następujących po sobie sekwencji NNCN (jeden z 4 oligonukleotydów NNCN_1, NNCN_2, NNCN_3 lub NNCN_4, Tabela 3.5.), rozpuszczano w jednokrotnie stężonym buforze reakcyjnym dla białka AID i denaturowano poprzez inkubację 5 min w temperaturze 95°C i nagłe schłodzenie na lodzie.
- Zdenaturowane ssDNA inkubowano z testowanym preparatem białkowym (w ilości 4,5 µg całkowitego białka) w jednokrotnie stężonym buforze reakcyjnym dla białka AID w temperaturze 37°C. Reakcję prowadzono w objętości 10 µl przez 30 minut.
- 1 µl mieszaniny reakcyjnej używano jako matrycy w reakcji PCR prowadzonej z wykorzystaniem gotowego zestawu odczynników KAPA HiFi HotStart Uracil+ ReadyMix firmy KAPA Biosystems. Skład mieszaniny reakcyjnej oraz warunki reakcji zawierają Tabele 3.2.16. oraz 3.2.17.
- Produkty reakcji PCR oczyszczano (podrozdział 3.2.1.2.), a następnie poddawano reakcji dodawania niesparowanych reszt adenozytowych do końców 3'.

Skład mieszaniny reakcyjnej:

- ✓ polimeraza DNA Taq (5 U/µl) – 0,5 µl
- ✓ bufor reakcyjny 10x (dostarczony wraz z polimerazą Taq) – 1 µl

✓ dATP (2,5 mM) – 1 μ l

✓ DNA – 7,5 μ l

Mieszaninę reakcyjną inkubowano 20 min w temperaturze 72°C.

- Produkty reakcji oczyszczano (podrozdział 3.2.1.2.), a następnie poddawano ligacji z wektorem pGEM-T Easy (podrozdział 3.2.1.9.).
- Produktem reakcji ligacji transformowano komórki kompetentne *E. coli* szczep One shot[®] TOP10 (podrozdział 3.2.3.3.). Selekcji komórek zawierających wektor z insertem dokonywano na podstawie testu α -komplementacji.
- Dwadzieścia pojedynczych kolonii bakteryjnych poddawano procedurze izolacji plazmidów w małej skali (podrozdział 3.2.1.10.), a następnie wyizolowane plazmidy sekwencjonowano (podrozdział 3.2.1.6.). Wyniki sekwencjonowania poddawano analizie bioinformatycznej. Obecność tranzykcji cytydyny do tymidyny świadczyła o zajściu reakcji deminacji.

Tabela 3.2.16. Skład mieszaniny reakcyjnej - PCR z wykorzystaniem gotowego zestawu odczynników KAPA HiFi HotStart Uracil+ ReadyMix firmy KAPA Biosystems.

odczynnik (stężenie początkowe)	ilość (stężenie końcowe)
matryca DNA	1 μ l
starter homologiczny do końca 5' powielanego fragmentu (pcrNNCN_F)	15 pmol
starter komplementarny do końca 3' powielanego fragmentu (pcrNNCN_R)	15 pmol
KAPA HiFi HotStart Uracil+ ReadyMix 2x	25 μ l
H ₂ O	do 50 μ l

Tabela 3.2.17. Warunki PCR z wykorzystaniem gotowego zestawu odczynników KAPA HiFi HotStart Uracil+ ReadyMix firmy KAPA Biosystems.

numer etapu	Etap	czas trwania	temperatura	liczba cykli
1	Denaturacja	2 min	95°C	-
2	Denaturacja	30 s	98°C	4
	Hybrydyzacja	30 s	57°C	
	Elongacja	2 min	72°C	
3	Denaturacja	30 s	98°C	26
	Hybrydyzacja	30 s	57°C	
	Elongacja	1 min	72°C	
4	Elongacja	1 min	72°C	-

3.2.4.6. Badanie wpływu AID na globalny poziom 5mC w genomowym DNA

Zmiany w globalnym poziomie 5mC w genomowym DNA (gDNA) w warunkach *in vitro* analizowano z wykorzystaniem gotowego zestawu odczynników 5-mC DNA ELISA Kit firmy Zymo Research zgodnie z instrukcją producenta. Jako matrycę DNA w reakcji wykorzystywano metylowany gDNA pochodzący z *E. coli* (wchodzące w skład w/w gotowego zestawu odczynników) lub metylowany, ludzki gDNA (dostępny komercyjnie, Zymo Research).

- 100 ng gDNA rozpuszczano w jednokrotnie stężonym buforze reakcyjnym dla białka AID. Próby denaturowano poprzez inkubację 5 min w temperaturze 95°C i nagłe schłodzenie na lodzie.
- Zdenaturowany gDNA inkubowano 8 lub 16 h z testowanym preparatem białkowym (w ilości 4,5 µg całkowitego białka) w temperaturze 37°C. Reakcję prowadzono w objętości 10 µl.
- Kolejne etapy reakcji obejmujące: denaturację badanych prób, opłaszczanie płytek wielodołkowych, inkubację w buforze blokującym, inkubację z przeciwciałami i detekcję kolorymetryczną, przeprowadzano zgodnie z instrukcją producenta.
- Dokonywano pomiaru absorpcji światła o długości fali 450 nm z wykorzystaniem czytnika płytek Victor™ X4 firmy Perkin Elmer.
- Jako kontrole negatywne stosowano próbki gDNA inkubowane z buforem reakcyjnym niezawierającym testowanego preparatu białkowego. Zarówno kontrole negatywne jak i próby ślepe (niezawierające gDNA) poddawano identycznej procedurze jak próby badane.
- Wykonywano trzy powtórzenia techniczne dla każdej z prób. Z wartości absorpcji dla trzech powtórzeń wyznaczano średnią arytmetyczną i odchylenie standardowe. Zmiany w globalnym poziomie metylacji gDNA prób badanych oceniano względem kontroli negatywnych.

3.2.5. Modelowanie struktury trzeciorzędowej AID, dynamika molekularna w polu siłowym *Charmm36*, dokowanie substratów do centrum aktywnego AID

Modelowanie struktury trzeciorzędowej AID

Model struktury trzeciorzędowej AID oraz jej mutantu N51A uzyskano dzięki uprzejmości profesora Janusza Bujnickiego z Międzynarodowego Instytutu Biologii Molekularnej i Komórkowej w Warszawie. Został on stworzony na drodze modelowania homologicznego z wykorzystaniem dwóch struktur krystalicznych domeny katalitycznej białka APOBEC3G (PDB 3e1u oraz 3v4k). Modele wygenerowano przy użyciu następujących narzędzi: (i) serwera SWISS-MODEL [178, 179]; (iii) metaservera GeneSilico [180] (do wygenerowania przyrównania sekwencji białek homologicznych (ang. *alignment*)); (ii) oprogramowania MODELLER [181] uruchomionego za pośrednictwem zestawu narzędzi bioinformatycznych GeneSilico TOOLKIT.

Dynamika molekularna w polu siłowym *Charmm36*

Dynamika molekularna otrzymanych modeli białek jak i modeli białek z zadokowanymi substratami została przeprowadzona w Zakładzie Chemii i Biologii Strukturalnej Kwasów Nukleinowych ICHB PAN przez dr Joannę Sarzyńską. Wszystkie symulacje zostały przeprowadzone przy użyciu oprogramowania NAMD w polu siłowym *Charmm36* metodą dynamiki Langevin [182, 183]. Dynamikę molekularną analizowano przy zastosowaniu następujących parametrów: temperatura 298 K, zespół izotermiczno-izobaryczny (stała liczba cząstek, stała temperatura i ciśnienie, NPT), współczynnik tłumienia 2/ps. Oddziaływania elektrostatyczne obliczano metodą Particle Mesh Ewald (PME). Układy równoważono (25 ns), a następnie prowadzono właściwe symulacje (50 lub 100 ns trajektorie) stosując ten sam zestaw parametrów.

Dokowanie substratów do centrum aktywnego AID i jej mutantów

Dokowanie substratów do centrum aktywnego modelu AID typu dzikiego oraz jej mutantu N51A zostało przeprowadzone w Zakładzie Biologii Molekularnej i Systemowej ICHB PAN przez dr. Karola Kamela. Substraty dokowano ręcznie (ang. *manual docking*) z uwzględnieniem mechanizmu reakcji deaminacji cytydyny (opis mechanizmu w podrozdziale 1.1.3.). Podczas pozycjonowania cytydyny

i 5-metylocytydyny uwzględniono następujące kryteria: (i) prawidłowe pozycjonowanie węgla C4 w stosunku do jonu cynku i aktywowanej cząsteczki wody (umożliwiające zajście ataku hydrofilowego na ten atom); (ii) bliskość grupy karboksylowej reszty E58 do fragmentu aminowego pierścienia cytydyny/5-metylocytydyny; (iii) w przypadku AID typu dzikiego, bliskość łańcucha bocznego reszty N51 i grupy karbonylowej cytydyny/5-metylocytydyny. Aby zapewnić obecność właściwych oddziaływań (niezbędnych do zajścia reakcji deaminacji) deprotonowano cząsteczkę wody w miejscu aktywnym enzymu (powstawał jon hydroksylowy) oraz protonowano grupę karboksylową reszty E58 (umożliwiało to powstanie wiązania wodorowego między resztą E58 a atomem azotu w pierścieniu cytydyny/5-metylocytydyny). W trakcie dokowania czteronukleotydowych motywów przyjęto kryterium pozycjonowania grup fosforanowych i zasad azotowych w taki sposób aby uniknąć odpychania ładunków elektrostatycznych o tym samym znaku. Po zadokowaniu substratu, struktury optymalizowano w próżni metodami mechaniki molekularnej stosując pole siłowe Universal Force Field51 [184].

4. WYNIKI

4.1. Otrzymywanie białka hAID w systemie prokariotycznym

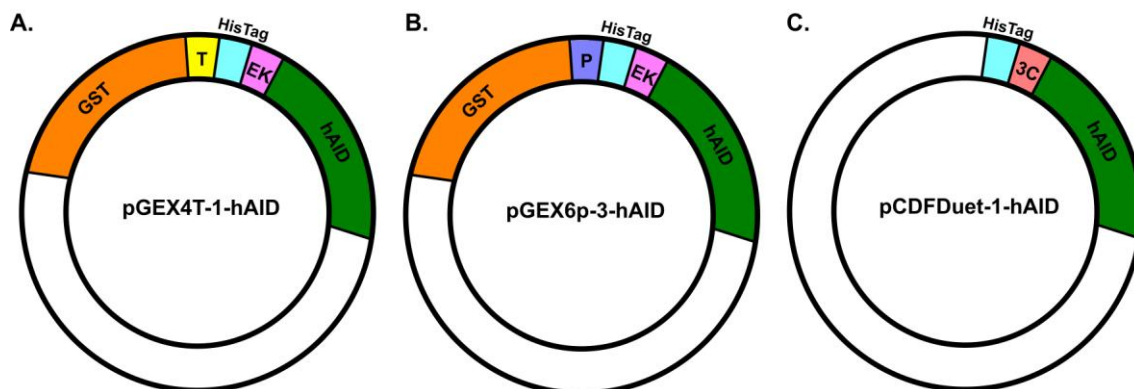
Przygotowanie konstruktów do ekspresji genu hAID w komórkach bakteryjnych.

Obecnie żaden preparat o aktywności AID nie jest dostępny handlowo. Dlatego w pierwszym etapie badań postawiono sobie za cel stworzenie wydajnego systemu produkcji hAID o pełnej aktywności enzymatycznej. Wybrano system bakteryjny, jako najbardziej wydajny z dostępnych systemów ekspresyjnych.

W pierwszej kolejności należało otrzymać konstrukt ekspresyjny do produkcji hAID. W tym celu posłużono się klonem cDNA hAID w wektorze pOTB7 (dostępnym komercyjnie w firmie Genecopoeia). Klon poddano sekwencjonowaniu z wykorzystaniem zestawu odczynników BigDye Terminator v1.1 Sequencing Kit (Applied Biosystems), zgodnie z instrukcją producenta (podrozdział 3.2.1.6). Otrzymaną sekwencję porównano z sekwencjami hAID, zdeponowanymi w internetowej bazie danych GenBank. W rezultacie potwierdzono zgodność sekwencji klonu w wektorze pOTB7 (pOTB7-hAID) z sekwencją zdeponowaną w bazie GenBank pod numerem NM_020661. Stwierdzono, że jest to gen typu dzikiego, nieobciążony mutacjami charakterystycznymi dla zespołu chorobowego HIGM2 czy też innymi zmianami. Wektor pOTB7-hAID posłużył następnie jako cząsteczka matrycowa w reakcji PCR, przeprowadzonej w celu namnożenia sekwencji kodującej hAID.

Na podstawie danych literaturowych można było sądzić, że hAID jest białkiem trudnym do produkcji i oczyszczenia. Aby zapewnić łatwiejszy przebieg tych procesów, do otrzymania hAID postanowiono wykorzystać wektor ekspresyjny pGEX4T-1 (GE Healthcare), zawierający sekwencję kodującą S-transferazę glutationową (GST). Spodziewano się, że zastosowanie tego znacznika powinno wyraźnie zwiększyć rozpuszczalność produkowanego białka i ułatwić jego oczyszczanie metodą chromatografii powinowactwa. Miejsce klonowania w plazmidzie pGEX4T-1 znajduje się poniżej sekwencji kodującej GST i krótkiego łącznika, kodującego sekwencję aminokwasową rozpoznawaną przez trombinę. Wprowadzenie insertu w tej samej ramce odczytu umożliwi produkcję białka fuzyjnego, a następnie oczyszczenie go na złożu zawierającym glutation. Dodatkowo w obrębie insertu wprowadzono sekwencję kodującą znacznik heksahistydynowy (HisTag) oraz sekwencję kodującą motyw rozpoznawany przez enterokinazę (Rysunek 4.1.1.A). Wybrana strategia umożliwia produkcję białka zaopatrzonego w znacznik GST i HisTag, a co za tym idzie, również

zastosowanie dwóch niezależnych metod oczyszczania: chromatografii powinowactwa do glutationu i chromatografii powinowactwa do jonów niklu. Ponadto miejsca rozpoznawane przez trombinę i enterokinazę powinny pozwolić na odcięcie obu znaczników i otrzymanie natywnego białka.



Rysunek 4.1.1 Konstrukty ekspresyjne przygotowane w celu produkcji hAID w systemie bakteryjnym. A) Konstrukt ekspresyjny przygotowany z wykorzystaniem wektora pGEX4T-1 zawierający: sekwencję kodującą GST (kolor pomarańczowy), sekwencję kodującą miejsce rozpoznawane przez trombinę (kolor żółty), sekwencję kodującą HisTag (kolor niebieski), sekwencję kodującą motyw rozpoznawany przez enterokinazę (kolor różowy) oraz sekwencję kodującą hAID (kolor zielony). B) Konstrukt ekspresyjny przygotowany z wykorzystaniem wektora pGEX6p-3. Konstrukt ten różni się od konstrukt pGEX4T-1-hAID sekwencją kodującą miejsce rozpoznawane przez proteazę – w tym przypadku jest to proteaza PreScission (kolor fioletowy). C) Konstrukt ekspresyjny przygotowany z wykorzystaniem wektora pCDFDuet-1 zawierający: sekwencję kodującą HisTag (kolor niebieski), sekwencję kodującą motyw rozpoznawany przez proteazę 3C (kolor czerwony) oraz sekwencję kodującą hAID (kolor zielony). Konstrukty A i B umożliwiały produkcję białka fuzyjnego o masie ok. 50 kDa. Konstrukt C umożliwiał produkcję białka fuzyjnego o masie ok. 25 kDa.

W celu uzyskania insertu kodującego hAID zaopatrzoną w sekwencję kodującą HisTag i sekwencję kodującą motyw rozpoznawany przez enterokinazę, przeprowadzono dwie reakcje PCR, z wykorzystaniem starterów: F_03_aid, F_04_aid, R_03_aid i wektora pOTB7-hAID, jako matrycy. Sekwencje starterów przedstawiono w Tabeli 3.1.5. Ze względu na znaczną długość dodatkowych sekwencji dołączanych do sekwencji kodującej hAID, amplifikację przeprowadzono w dwóch etapach. W pierwszym posłużono się starterem F_03_aid, homologicznym do końca 5' namnażanego fragmentu genu, który wprowadzał sekwencję kodującą motyw rozpoznawany przez enterokinazę oraz starterem R_03_aid, komplementarnym do końca 3' namnażanego fragmentu genu. Starter R_03_aid wprowadzał miejsce rozpoznawane przez endonukleazę EcoRI i dwa kodony STOP. Produkt PCR

rozcieńczono stukrotnie i wykorzystano jako matrycę w drugiej reakcji PCR. Posłużono się w niej starterem F_04_aid, który wprowadzał sekwencję kodującą HisTag oraz miejsce rozpoznawane przez endonukleazę BamHI. Jako drugiego startera użyto ponownie R_03_aid. Obie reakcje PCR przeprowadzono z wykorzystaniem polimerazy DNA Pfu o wysokiej dokładności kopiowania (Promega) (podrozdział 3.2.1.1.). Produkty reakcji PCR analizowano metodą elektroforezy w żelu agarozowym w warunkach natywnych (podrozdział 3.2.1.3.). W kolejnym etapie produkty drugiej reakcji PCR oraz plazmid pGEX-4T-1 cięto enzymami restrykcyjnymi BamHI i EcoRI (podrozdział 3.2.1.7.). Produkty cięcia analizowano metodą elektroforezy w żelu agarozowym w warunkach natywnych. Tak przygotowany plazmid i insert poddano ligacji z użyciem ligazy DNA T4 (Fermentas) (podrozdział 3.2.1.8.). Produktami ligacji transformowano komórki kompetentne *E. coli* DH5 α (podrozdział 3.2.3.3.). Transformowane bakterie hodowano na szalkach ze stałą pożywką LB (ang. *Luria Broth*) zawierającą ampicylinę (stężenie końcowe 200 $\mu\text{g/ml}$). Uzyskane pojedyncze kolonie wykorzystano do przygotowania płynnych kultur w pożywce LB zawierającej ampicylinę, z których następnie izolowano plazmidy metodą lizy alkalicznej (podrozdział 3.2.1.10.). Prawidłowość procesu klonowania potwierdzono tnąc wyizolowane plazmidy enzymami BamHI i EcoRI oraz analizując uzyskane produkty metodą elektroforezy w żelu agarozowym w warunkach natywnych. W wyniku cięcia plazmidów zawierających insert wyzwalał był dsDNA o długości ok. 630 par zasad. Plazmidy wyzwalające fragmenty o oczekiwanej długości sekwencjonowano. Na tej podstawie dokonano wyboru prawidłowego konstruktu pGEX4T-1-hAID, kodującego hAID zaopatrzoną w znacznik GST, HisTag oraz motywy rozpoznawane przez proteazy.

Równocześnie przygotowano dwa alternatywne konstrukty: (i) analogiczny konstrukt jak opisany powyżej na bazie wektora pGEX6p-3 (GE Healthcare) (Rysunek 4.1.1.B); (ii) plazmid pCDFDuet-1 (Novagen) z wklonowaną sekwencją kodującą hAID (konstrukt otrzymany dzięki uprzejmości dr Niny Papavasiliou z The Rockefeller University (Nowy Jork, USA)) (Rysunek 4.1.1.C). Miejsce klonowania w plazmidzie pGEX6p-3 znajduje się poniżej sekwencji kodującej GST i krótkiego łącznika, kodującego sekwencję aminokwasową rozpoznawaną przez proteazę PreScission. Procedura klonowania do wektora pGEX6p-3 przebiegała identycznie jak w przypadku wektora pGEX4T-1. Konstrukty pGEX6p-3-hAID oraz pGEX4T-1-hAID różniły się zatem sekwencjami kodującymi miejsca rozpoznawane przez odpowiednie proteazy.

Otrzymany dzięki uprzejmości dr Niny Papavasiliou konstrukt pCDFDuet-1-hAID umożliwił produkcję białka fuzyjnego hAID ze znacznikiem HisTag i miejscem rozpoznawanym przez proteazę 3C. Sekwencja kodująca miejsce rozpoznawane przez proteazę była wprowadzona jako element insertu, natomiast sekwencja kodująca HisTag znajdowała się w wektorze. Otrzymany konstrukt sekwencjonowano i na podstawie uzyskanych wyników potwierdzono jego poprawność.

Produkcja białka hAID w komórkach bakteryjnych. W celu przeprowadzenia ekspresji genu kodującego hAID w komórkach bakteryjnych, otrzymany wektor ekspresyjny pGEX4T-1-hAID wprowadzono do komórek kompetentnych *E. coli* na drodze transformacji metodą szoku cieplnego (podrozdział 3.2.3.3.). Ponieważ wydajność produkcji białka może zależeć od zastosowanego szczepu *E. coli*, testowano następujące komórki: BL21(DE3)pLysS (Novagen), Origami(DE3)pLysS (Novagen), BL21-ALTM One Shot[®] (Invitrogen), Rosetta(DE3)pLysS (Novagen). Krórką charakterystykę testowanych szczepów przedstawiono w Tabeli 3.1.6. Spośród otrzymanych na szalce kolonii bakteryjnych transformowanych wektorem pGEX4T-1-hAID, wybrano jedną i sporządzono płynną hodowlę. Hodowla ta posłużyła następnie do zaszczepienia większej ilości pożywki i produkcji białka (podrozdział 3.2.3.5.).

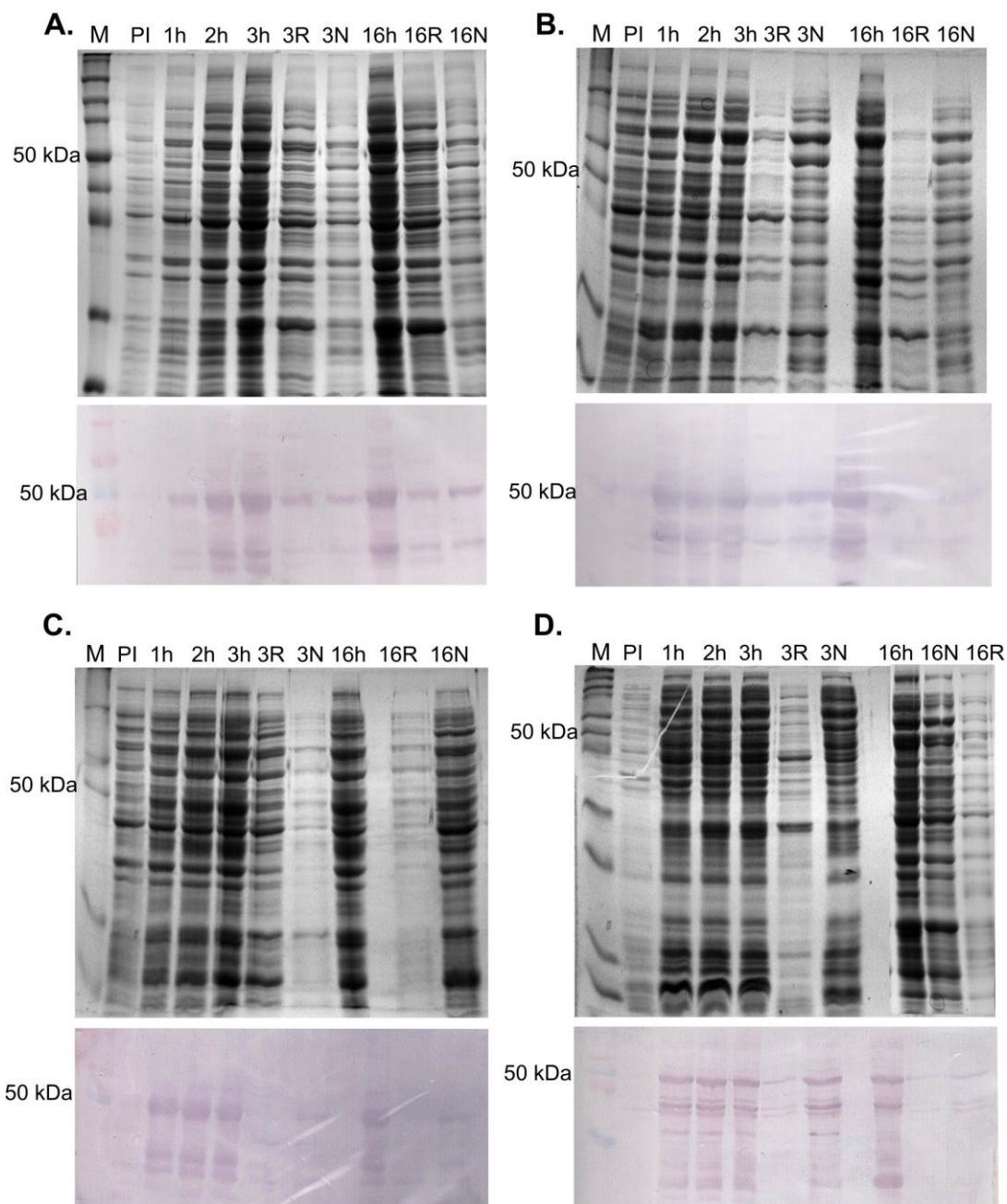
Aby ustalić optymalne warunki hodowli oraz porównać wydajność produkcji białka hAID¹⁴ w testowanych szczepach bakterii, pierwszą serię wstępnych eksperymentów przeprowadzono w skali analitycznej (podrozdział 3.2.3.5.). Zaszczepione hodowle o objętości 20 ml pożywki LB zawierające odpowiedni antybiotyk (Tabela 3.2.13) inkubowano w temperaturze 37°C wytrząsając z prędkością 250 rpm. Hodowle bakterii prowadzono w tych samych warunkach do chwili osiągnięcia przez nie gęstości optycznej $OD_{600} = 0,6 - 0,8$. Wówczas zachodzą najbardziej intensywne podziały komórek (faza wzrostu logarytmicznego), jest to zatem najkorzystniejszy moment by indukować ekspresję heterologicznego genu. Tuż przed indukcją, z niewielkiej ilości hodowli sporządzano stoki glicerolowe (podrozdział 3.2.3.5.), które można było przechowywać przez dłuższy czas w temperaturze -80°C. Stoki te wykorzystywano w kolejnych eksperymentach, podczas produkcji białka w skali preparatywnej. W pozostałej części hodowli bakteryjnej indukowano ekspresję heterologicznego genu, wklonowanego poniżej promotora *tac*, dodając do pożywki

¹⁴ W dalszej części pracy określenie „hAID” w odniesieniu do białka oznacza, w domyśle, rekombinowane białko fuzyjne.

izopropyl- β -D-1-tiogalaktopiranozydu (IPTG) do uzyskania stężenia końcowego 0,5 mM. Promotor *tac* (będący hybrydą promotorów *lac* i *trp*) jest wyłączany przez represor laktozowy (białko LacI, sekwencja kodująca to białko znajduje się w wektorze). IPTG wiąże się do represora laktozowego, powodując jego inaktywację. Umożliwia to inicjację transkrypcji wklonowanego do plazmidu genu i produkcję białka hAID. W przypadku szczepu bakterii BL21-ALTMOneShot[®] produkcję rekombinowanego białka hAID indukowano poprzez dodanie do pożywki IPTG do uzyskania stężenia końcowego 1 mM.

Poziom białka hAID analizowano przed indukcją ekspresji oraz po indukcji w trzech punktach czasowych – po 1, 2 i 3 h (hodowla prowadzona w temperaturze 37°C) lub alternatywnie w jednym punkcie czasowym – po 16 h (hodowla prowadzona w temperaturze 18°C). Analizy dokonywano metodą elektroforezy w żelu poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.2.1.). Ponieważ w każdym z testowanych szczepów bakteryjnych białko hAID było produkowane na stosunkowo niskim poziomie, jego identyfikacja była utrudniona i wymagała potwierdzenia metodą *western blot* (podrozdział 3.2.2.2.). W metodzie tej jako przeciwciała I-rzędowe wykorzystano komercyjnie dostępne kozie poliklonalne przeciwciała przeciwko hAID (Sigma). Detekcję przeprowadzono metodą kolorymetryczną – przeciwciała II-rzędowe (królicze poliklonalne przeciwciała przeciwko kozim IgG (ABCAM)) były sprzężone z fosfatazą alkaliczną. W przypadku ekstraktu białkowego pobranego po 3 h lub 16 h analizowano również rozpuszczalność białek (podrozdział 3.2.3.6.)

Wynik analizy ekstraktów białkowych izolowanych z czterech testowanych szczepów bakterii (BL21(DE3)pLysS, Origami(DE3)pLysS, BL21-ALTM One Shot[®], Rosetta(DE3)pLysS) transformowanych wektorem pGEX4T-1-hAID przedstawiono na Rysunku 4.1.2.A-D.



Rysunek 4.1.2. Analiza preparatów białkowych izolowanych z czterech szczepów bakteryjnych: BL21(DE3)pLysS (A), BL21-AL™ One Shot® (B), Origami(DE3)pLysS (C), Rosetta(DE3)pLysS (D), transformowanych wektorem pGEX-4T-1-hAID. Dla każdego szczepu przeprowadzono analizę metodą elektroforezy w 12 lub 15% żelu poliakrylamidowym, w warunkach denaturujących oraz metodą *western blot*. Oczekiwana masa rekombinowanego białka hAID wynosiła 50 kDa. M – wzorzec mas; PI – ekstrakty białkowe izolowane z bakterii przed indukcją ekspresji heterologicznego genu; 1h/2h/3h – ekstrakty białkowe izolowane z bakterii 1, 2 lub 3 h po indukcji ekspresji heterologicznego genu (bakterie hodowane w temperaturze 37°C); 3R/3N – frakcje białek rozpuszczalnych / nierozpuszczalnych izolowane z bakterii 3 h po indukcji ekspresji heterologicznego genu; 16 h - ekstrakty białkowe izolowane z bakterii 16 h po indukcji ekspresji heterologicznego genu (bakterie hodowane w temperaturze 18°C); 16R/16N - frakcje białek rozpuszczalnych / nierozpuszczalnych izolowane z bakterii 16 h po indukcji ekspresji heterologicznego genu.

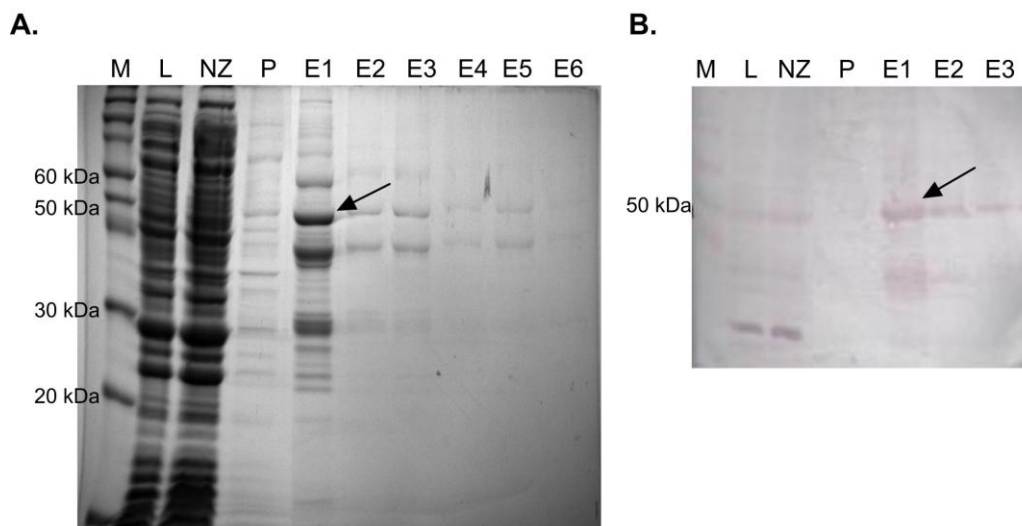
We wszystkich badanych ekstraktach stwierdzono obecność białka hAID. Najbardziej wydajną produkcję stwierdzono w szczepie bakterii BL21(DE3)pLysS. Jednocześnie znaczna ilość białka hAID otrzymanego w tych bakteriach była obecna we frakcji białek rozpuszczalnych. Izolacja i oczyszczanie hAID z frakcji białek nierozpuszczalnych wiązałyby się z koniecznością przeprowadzenia denaturacji i renaturacji białka. Procesy te byłyby jednak niepożądane ze względu na ich potencjalny negatywny wpływ na strukturę otrzymywanego białka i jego aktywność. Dlatego też zdecydowano się izolować i oczyszczać hAID z frakcji białek rozpuszczalnych. Porównanie wydajności produkcji rozpuszczalnego białka w hodowli prowadzonej 3 h w 37°C oraz w hodowli prowadzonej 16 h w 18°C wskazało, że drugi z zastosowanych wariantów jest efektywniejszy. Obniżona temperatura hodowli bakteryjnej ograniczała powstawanie ciałek inkluzyjnych co jest szczególnie istotne w przypadku białek potencjalnie szkodliwych dla komórki (takich jak hAID).

Do wybranego w wyniku powyższych analiz szczepu BL21(DE3)pLysS wprowadzono na drodze transformacji metodą szoku cieplnego (podrozdział 3.2.3.3.) również alternatywne konstrukty ekspresyjne: (i) pGEX6p-3-hAID lub (ii) pCDFDuet-1-hAID. Bakterie transformowane wektorem pGEX6p-3-hAID lub pCDFDuet-1-hAID użyto następnie do produkcji hAID w skali analitycznej w sposób analogiczny jak opisany powyżej dla konstruktu pGEX4T-1-hAID. W przypadku wektora pCDFDuet-1-hAID, gen heterologiczny jest wklonowywany poniżej promotora T7. Dodanie IPTG do hodowli, indukuje ekspresję genu polimerazy RNA bakteriofaga T7 (zintegrowanego z genomem bakterii). Enzym ten rozpoznaje promotor T7, co umożliwia ekspresję heterologicznego genu i produkcję białka. Doświadczenia w skali analitycznej potwierdziły obecność białka hAID w badanych ekstraktach białkowych jak i jego obecność we frakcjach białek rozpuszczalnych. Różnice w profilach ekspresji *hAID* pomiędzy poszczególnymi układami ekspresyjnymi były niewielkie. Ponieważ spodziewano się znaczących różnic na etapie oczyszczania białka, do produkcji hAID w skali preparatywnej zostały użyte wszystkie trzy konstrukty.

Izolacja i oczyszczanie hAID. Produkcję hAID w skali preparatywnej (podrozdział 3.2.3.7.) prowadzono w komórkach BL21(DE3)pLysS zawierających jeden z konstruktów ekspresyjnych: pGEX4T-1-hAID, pGEX6p-3-hAID lub pCDFDuet-1-hAID. Z hodowli o objętości 1 l prowadzonej przez 16 h w temperaturze 18°C (od

momentu indukcji ekspresji) izolowano białka. W tym celu zawiesinę bakterii wirowano, a osad zawieszano w buforze odpowiednim dla danej metody oczyszczania (podrozdział 3.2.3.8.).

W celu uzyskania jak największej homogenności preparatu białkowego, podczas procedury oczyszczania hAID zastosowano kilka podejść. Pierwszą metodą oczyszczania hAID z bakterii transformowanych konstruktem pGEX4T-1-hAID była chromatografia powinowactwa do glutationu (podrozdział 3.2.3.8.). W tym celu osad komórek bakteryjnych zawieszono w 50 ml buforu 1x PBS zawierającego koktajl inhibitorów proteaz (Roche Diagnostics), ditiotreitól (DTT) w stężeniu końcowym 1 mM, oraz Triton X-100 w stężeniu końcowym 1%. Dodatek inhibitorów proteaz zapobiega degradacji białek, natomiast zastosowanie słabego detergentu, jakim jest Triton X-100, wspomaga uwalnianie białek z błon cytoplazmatycznych. Zdolności redukcyjne DTT są wykorzystywane w celu uniemożliwienia tworzenia międzycząsteczkowych mostków disiarczkowych pomiędzy resztami cystein w białkach. Tak powstałą zawiesinę bakterii inkubowano w temperaturze 4°C przez 30 minut, a następnie wirowano w celu oddzielenia frakcji białek rozpuszczalnych od nierozpuszczalnych. Białko hAID obecne było w supernatancie, co potwierdziły analizy w żelach poliakrylamidowych w warunkach denaturujących oraz analizy metodą *western blot*. Supernatant zawierający frakcję białek rozpuszczalnych wykorzystano następnie do wstępnego oczyszczenia produkowanego białka na złożu glutationowym Glutathione Sepharose 4 Fast Flow (GE Healthcare). Glutation jest specyficznym ligandem dla GST, stąd unieruchomiony na stałym podłożu umożliwia oczyszczanie białek znakowanych GST. Podczas oczyszczania białka zastosowano procedurę zgodną z instrukcją producenta. Po odmyciu białek nie wiążących się z glutationem frakcję białek związanych ze złożem eluowano przemywając kolumnę buforem 50 mM TrisHCl, pH = 8, zawierającym zredukowany glutation o stężeniu końcowym 10 mM. Nadmiar wolnego glutationu uwalnia związane ze złożem białko fuzyjne. Zbierane frakcje analizowano metodą elektroforezy w żelu poliakrylamidowym oraz techniką *western blot*. Wyniki obu analiz przedstawia Rysunek 4.1.3. Zastosowana metoda oczyszczania pozwoliła usunąć większość białek bakteryjnych. Uzyskany w jej wyniku preparat nie był jednak homogeny, a zawartość hAID w tym preparacie wynosiła maksymalnie 30%.

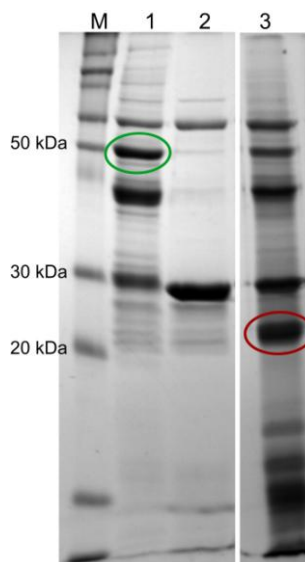


Rysunek 4.1.3. Analiza frakcji otrzymanych w wyniku oczyszczania hAID na złożu glutationowym.

A) Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących następujących frakcji: L – lizat bakteryjny inkubowany ze złożem; NZ – frakcja białek niezwiązanych ze złożem; P – frakcja białek uzyskana po przemyciu złoża buforem niezawierającym glutationu, od E1 do E6 – frakcje eluowane ze złoża buforem zawierającym glutation; M – wzorzec mas; strzałką oznaczono prążek odpowiadający białku hAID. B) analiza wybranych frakcji metodą *western blot* (oznaczenia analogiczne jak w punkcie A).

Wstępnie oczyszczone na kolumnie wypełnionej złożem glutationowym białko hAID zawierało w swej sekwencji aminokwasowej miejsce rozpoznawane przez enterokinazę. Spodziewano się, że zastosowanie tego enzymu proteolitycznego pozwoli na odcięcie znaczników: GST oraz HisTag i otrzymanie natywnego białka. Ponowne zastosowanie chromatografii powinowactwa do glutationu wobec białka pozbawionego GST powinno skutkować zatem otrzymaniem preparatu o znacznie większej homogenności. Aby zrealizować powyższą strategię oczyszczania, wymieniono bufor w preparacie białkowym wstępnie oczyszczonym na kolumnie glutationowej na bufor reakcyjny dla rekombinowanej enterokinazy i poddano trawieniu proteolitycznemu z zastosowaniem gotowego zestawu odczynników: Enterokinase Cleavage Capture Kit (Novagen), zgodnie z instrukcją producenta. Reakcję cięcia prowadzono przez 16 h w temperaturze pokojowej. Enterokinazę usuwano następnie z mieszaniny reakcyjnej przy pomocy złoża (EKapture Agarose) i filtrów dostarczonych przez producenta enzymu. Pomimo prób optymalizacji cięcia, w następstwie tej reakcji dochodziło do wytrącenia białek. Produkty reakcji wirowano, a osad i supernatant analizowano w żelu poliakrylamidowym w warunkach denaturujących. Analizy wykazały, że rekombinowane białko było cięte z wydajnością bliską 100%, ale po odcięciu

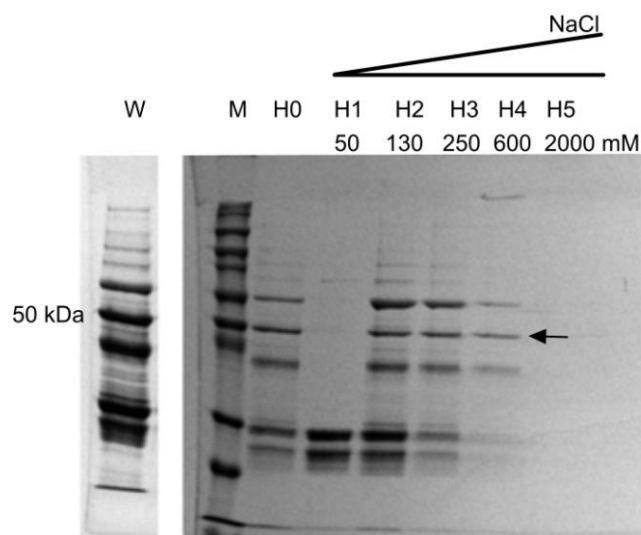
znaczników zdecydowana większość hAID wytrącała się (Rysunek 4.1.4.). Wraz z hAID wytrąceniu ulegały również białka bakteryjne obecne w preparacie. Na tej podstawie stwierdzono, że obecność znacznika GST drastycznie zwiększa rozpuszczalność produkowanego białka i zapewnia jego stabilność w roztworze. Zaniechano zatem dalszych prób jego usuwania.



Rysunek 4.1.4. Analiza metodą elektroforezy w 12% żelu poliakrylamidowym w warunkach denaturujących preparatów białkowych uzyskanych w wyniku cięcia proteolitycznego enterokinazą. M – wzorzec mas; 1 – preparat białkowy przed cięciem proteolitycznym, 2 – preparat białkowy po cięciu enterokinazą, frakcja obecna w supernatancie; 3 – preparat białkowy po cięciu enterokinazą, frakcja obecna w osadzie; zielonym okręgiem zaznaczono białko fuzyjne hAID przed cięciem proteolitycznym; czerwonym okręgiem zaznaczono hAID pozbawione GST (opis w tekście).

Ponieważ preparat białkowy uzyskany w wyniku oczyszczania na kolumnie wypełnionej złożem glutationowym nie był homogenny, poszukiwano dalszych sposobów jego oczyszczenia. Zastosowano zatem kolumnę wypełnioną złożem heparynowym: Heparin Sepharose 6 Fast Flow (GE Healthcare) (podrozdział 3.2.3.8.). Złoże heparynowe daje możliwość oczyszczenia białek na zasadzie chromatografii powinowactwa do heparyny. Powinowactwo to jest charakterystyczne dla wielu białek wiążących DNA. Preparat białkowy, powstały poprzez połączenie frakcji eluowanych z kolumny wypełnionej złożem glutationowym, łączono z 2 ml złoża heparynowego, inkubowano 30 min w temperaturze pokojowej, a następnie zawiesinę nakładano na kolumnę chromatograficzną. Podczas oczyszczania białka zastosowano procedurę zgodną z instrukcją producenta. Frakcję białek związanych ze złożem eluowano przemywając kolumnę gradientem roztworu NaCl (stosowano stężenia: 50, 130, 250,

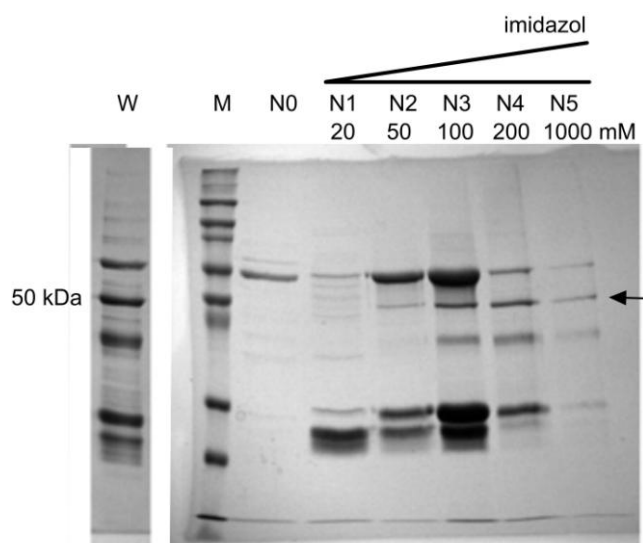
600, 2000 mM) w 50 mM Tris-HCl, pH = 8. Zebrane frakcje zagęszczano (podrozdział 3.2.3.9.) i analizowano metodą elektroforezy w żelach poliakrylamidowych w warunkach denaturujących. Wynik analizy przedstawia Rysunek 4.1.5. Stwierdzono obecność hAID w większości analizowanych frakcji. Oczyszczanie tą metodą nie przyniosło poprawy homogenności preparatu.



Rysunek 4.1.5. Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących frakcji otrzymanych w wyniku oczyszczania hAID na złożu heparynowym. W – preparat białkowy poddany wiązaniu ze złożem; M – wzorzec mas; H0 – frakcja białek niezwiązanych ze złożem; od H1 do H5 – frakcje białek eluowanych ze złoża gradientem roztworu NaCl (pod nazwą frakcji podano stężenie NaCl); strzałką oznaczono prążek odpowiadający białku hAID.

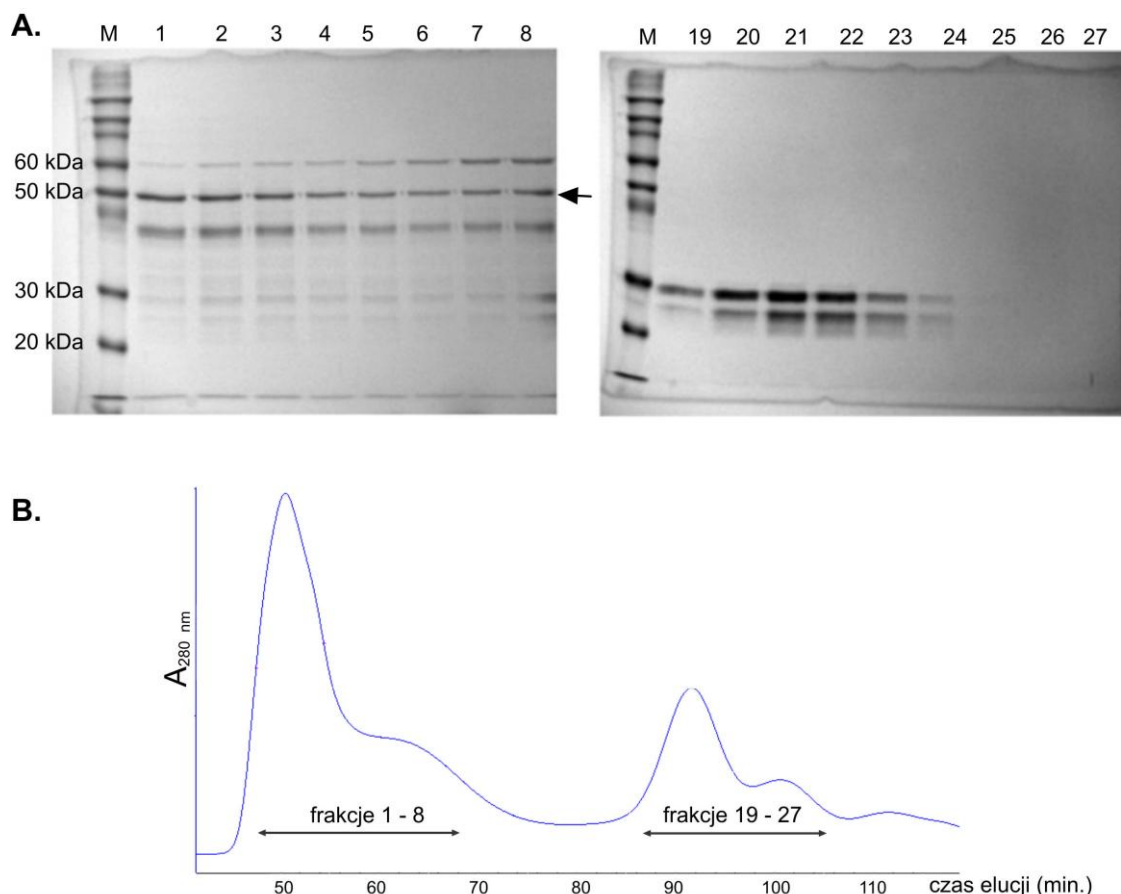
W kolejnym podejściu preparat białkowy wstępnie oczyszczony na kolumnie wypełnionej złożem glutationowym, poddano oczyszczaniu wykorzystując chromatografię powinowactwa do jonów niklu (podrozdział 3.2.3.8.). W preparacie białkowym, powstałym poprzez połączenie frakcji eluowanych z kolumny wypełnionej złożem glutationowym, wymieniano bufor na bufor I. Następnie dodawano jedną objętość buforu II i inkubowano ze złożem niklowym: Ni-NTA Agarose (Qiagen). Zawiesinę nakładano na kolumnę chromatograficzną. Kolumnę przemywano gradientem imidazolu (stosowano stężenia: 20, 50, 100, 200 mM) w jednokrotnie stężonym buforze III. Zebrane frakcje o objętości 2,5 ml zagęszczano do objętości 250 μ l i analizowano metodą elektroforezy w żelu poliakrylamidowych w warunkach denaturujących. W wyniku analizy (Rysunek 4.1.6.) stwierdzono obecność hAID we frakcjach wmywanych imidazolem o stężeniu od 50 mM do 1M. Frakcja 100 mM (N3) w porównaniu z preparatem wyjściowym zawierała znacznie gorszy stosunek

ilości białka hAID do białka bakteryjnego o masie 60 kDa oraz dwóch innych białek o masie 20 - 30 kDa. Frakcja 200 mM (N4) również nie była homogenna ale jakość preparatu uległa nieznacznej poprawie.



Rysunek 4.1.6. Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących frakcji otrzymanych w wyniku oczyszczania hAID na złożu niklowym. W – preparat białkowy poddany wiązaniu ze złożem; M – wzorzec mas; N0 – frakcja białek niezwiązanych ze złożem; od N1 do N5 – frakcje białek eluowanych ze złoża gradientem imidazolu (pod nazwą frakcji podano stężenie imidazolu); strzałką oznaczono prążek odpowiadający białku hAID.

W innym podejściu, preparat białkowy wstępnie oczyszczony na kolumnie wypełnionej złożem glutationowym, poddano wymianie buforu a następnie oczyszczano metodą sączenia molekularnego na kolumnie Superdex75 16/60 HiLoad firmy GE Healthcare. Sączenie molekularne przeprowadzono dwukrotnie w dwóch alternatywnych roztworach: (i) 50 mM Tris-HCl, pH = 8, 25 mM NaCl, 3% glicerol, 1 mM DTT; (ii) 50 mM Tris-HCl, pH = 8, 200 mM NaCl, 3% glicerol, 0,01% DDM, 2mM TCEP. Analiza frakcji zebranych po sączeniu molekularnym w pierwszym roztworze (Rysunek 4.1.7.) potwierdziła obecność hAID w kilku początkowych frakcjach. We frakcjach tych było również obecne kilka białek bakteryjnych o masie od 20 do 60 kDa. Świadczyło to o tym, że białka te oddziałując ze sobą oraz z hAID tworząc różnej wielkości multimery lub też agregaty. W drugim wariantcie zastosowano zatem znacznie wyższe stężenie soli (NaCl), silniejszy i stabilniejszy czynnik redukujący (TCEP) oraz niejonowy detergent (DDM). Zastosowanie drugiego buforu skutkowało niemalże identycznymi efektami oczyszczania.



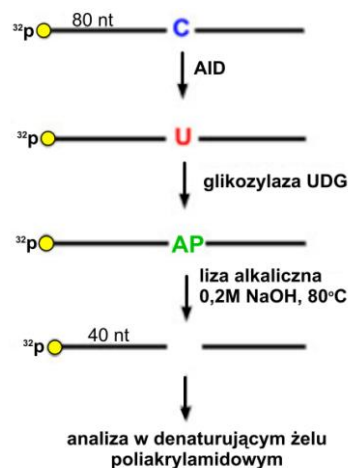
Rysunek 4.1.7. Analiza frakcji otrzymanych w wyniku oczyszczania hAID metodą sączenia molekularnego. A) Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących frakcji od 1 do 8 oraz od 19 do 27. M – wzorzec mas; strzałką oznaczono prążek odpowiadający białku hAID. B) Wykres wygenerowany przez kolektor frakcji, przedstawiający wartości stężeń całkowitego białka w poszczególnych frakcjach (pomiar spektrofotometryczny). Na jego podstawie dokonano wyboru frakcji, które poddano analizie metodą elektroforezy.

Wstępne testy aktywności deaminazowej hAID. Opisanie powyżej metody oczyszczania miały na celu uzyskanie preparatu o możliwie dużej homogenności przy jednoczesnym zachowaniu właściwości enzymatycznych produkowanego białka. Po każdym z wymienionych wyżej etapów oczyszczania dokonywano zatem oceny aktywności deaminazowej preparatu białkowego zawierającego hAID.

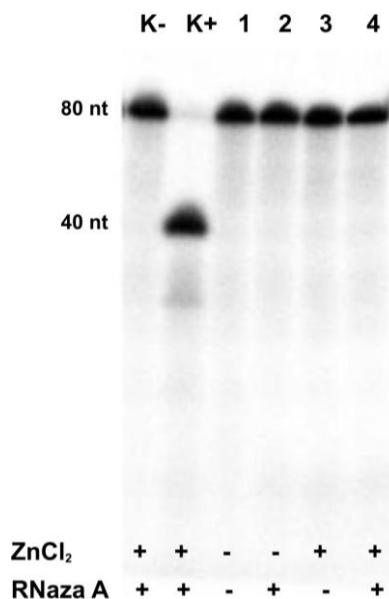
W celu wykazania, że produkowane białko wykazuje aktywność deaminazową wobec pojedynczej cytydyny znajdującej się w obrębie czteronukleotydowego motywu WRCY¹⁵ (będącego preferowanym przez hAID motywem) stosowano metodę

¹⁵ W = A/T; R = A/G; Y = C/T

opublikowaną w 2003 roku przez Sohail i współautorów [176] (Rysunek 4.1.8.) (podrozdział 3.2.4.2.). Jako substratu do reakcji używano 80-nukleotydowej cząsteczki ssDNA (Oligo_C – zawierającej pojedynczą resztę C w pozycji 40, Tabela 3.5.) znakowanej radioizotopowo na końcu 5' (podrozdział 3.2.4.1.). Oligomer ten inkubowano 30 minut w temperaturze 37°C z testowanym preparatem białkowym (w ilości 4,5 µg całkowitego białka). W przypadku obecności aktywnej deaminazy cytydyny, dochodzi do deaminacji cytydyny do urydyny. W kolejnym etapie do reakcji dodawano glikozylazę UDG (Fermentas) i inkubowano dalsze 30 minut w temperaturze 37°C. Enzym ten usuwa uracyl z DNA, pozostawiając miejsce apirymidynowe (AP). Następnie przeprowadzano hydrolizę alkaliczną wiązania fosfodiesterowego w miejscu AP. Produkty reakcji analizowano metodą elektroforezy w 15% żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.1.4.). W przypadku obecności w testowanym preparacie aktywnej deaminazy cytydyny, spodziewano się zaobserwować produkt o długości 40 nukleotydów. Równocześnie przeprowadzono reakcje kontrolne: (i) pozytywną reakcję kontrolną, bez preparatu białkowego, ale z wykorzystaniem cząsteczki substratowej posiadającej U zamiast C w pozycji 40 (Oligo_U, Tabela 3.5); (ii) negatywną reakcją kontrolną, bez preparatu białkowego, ale z wykorzystaniem standardowej cząsteczki substratowej (Oligo_C).



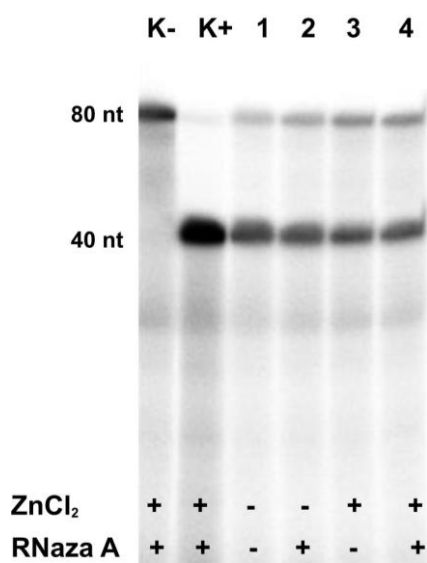
Rysunek 4.1.8. Schemat ilustrujący metodę badania aktywności deaminazowej hAID wobec pojedynczego motywu WRCY (opis w tekście).



Rysunek 4.1.9. Przykładowy wynik analizy metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących produktów powstających w trój etapowym teście aktywności deaminazowej preparatów białkowych. Aktywność badanego preparatu białkowego testowano w obecności chlorku cynku lub/i RNazy A (ścieżki 1 – 4). „K-” – negatywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca C w pozycji 40); „K+” – pozytywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca U w pozycji 40). Widoczny jest brak aktywności deaminazowej badanego preparatu.

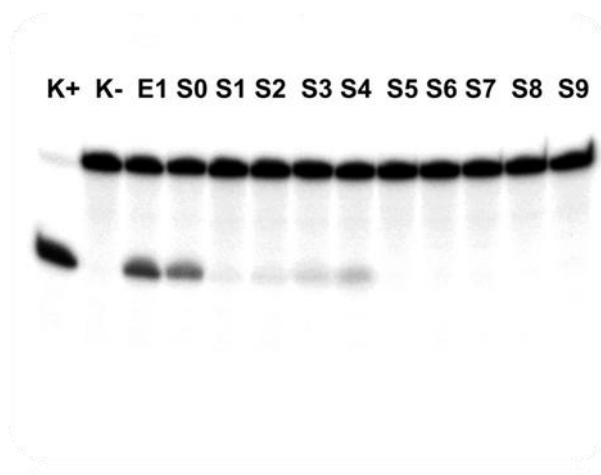
Początkowo żaden z testowanych preparatów białkowych nie wykazywał aktywności deaminazowej. W pierwszej kolejności próbowano zatem optymalizować warunki prowadzonych reakcji. Dane literaturowe sugerują, że związany z AID bakteryjny RNA może hamować jej aktywność deaminazową [32]. Dlatego podczas badania aktywności preparatów testowano wpływ RNazy A na przebieg reakcji. Wiadomo również, że białko AID ma w swej cząsteczce skoordynowany jon cynku, dlatego testowano przebieg reakcji w obecności chlorku cynku. Szczegółowy skład poszczególnych mieszanin reakcyjnych zawiera Tabela 3.2.15. Przeprowadzone testy aktywności wybranych preparatów białkowych wykazały brak aktywności deaminazowej bez względu na obecność w mieszaninie reakcyjnej RNazy A czy też jonów cynku (Rysunek 4.1.9.). Ponownie skupiono się zatem na optymalizacji produkcji i izolacji rekombinowanego białka. Między innymi testowano aktywność preparatu uzyskanego w wyniku oczyszczania na złożu glutationowym ekstraktu bakterii transformowanych konstruktem pGEX6p-3-hAID czy też preparatu uzyskanego w wyniku oczyszczania na złożu niklowym ekstraktu bakterii transformowanych

konstruktem pCDFDuet-1-hAID. Zastosowanie tych alternatywnych konstruktów nie przyniosło jednak oczekiwanego efektu. Po wielu nieudanych próbach okazało się, że niezbędna do otrzymania aktywnej formy białka jest suplementacja pożywki LB, w której prowadzona jest hodowla, chlorkiem cynku w stężeniu końcowym 0,06 mM. Atom cynku jest skoordynowany w centrum katalitycznym AID i jest niezbędny do jej aktywności. Zawartość tego pierwiastka w podstawowej pożywce LB nie była wystarczająca, co skutkowało obserwowanym brakiem aktywności testowanych preparatów białkowych. Profile ekspresji jak i jakość preparatów białkowych otrzymywanych w wyniku oczyszczania lizatu bakterii transformowanych konstruktem pGEX4T-1-hAID nie różniły w przypadku hodowli prowadzonej z dodatkiem chlorku cynku jak i bez jego dodatku. Zastosowane stężenie $ZnCl_2$ było na tyle niskie, że nie zakłócało procesów metabolicznych bakterii, a produkowane preparaty białkowe różniły się jedynie obserwowaną aktywnością deaminazową.



Rysunek 4.1.10. Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących produktów powstających w trójetapowym teście aktywności deaminazowej preparatu białkowego uzyskanego w wyniku oczyszczania lizatu bakterii transformowanych konstruktem pGEX4T-1-hAID, hodowanych w pożywce z dodatkiem chlorku cynku. Aktywność badanego preparatu białkowego testowano w obecności $ZnCl_2$ lub/i RNazy A (ścieżki 1 – 4). „K-” – negatywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca C w pozycji 40); „K+” – pozytywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca U w pozycji 40). Widoczna jest aktywność deaminazowa badanego preparatu, bez względu na obecność w mieszaninie reakcyjnej $ZnCl_2$ czy RNazy A.

Jak już wspomniano, celem podjętych działań było otrzymanie preparatu maksymalnie homogennego, posiadającego pełną aktywność enzymatyczną. Jak się jednak okazało, konieczne było wypracowanie rozwiązania kompromisowego. W wyniku przeprowadzanych analiz aktywności deaminazowej hAID na różnych etapach oczyszczania stwierdzono, że najwyższą aktywnością cechuje się preparat pochodzący z bakterii transformowanych konstruktem pGEX4T-1-hAID, wstępnie oczyszczony na kolumnie wypełnionej złożem glutationowym (Rysunek 4.1.10). Na wynik reakcji nie miała jednak wpływu obecność w mieszaninie reakcyjnej RNazy czy też $ZnCl_2$. Zastosowanie dodatkowych metod oczyszczania skutkowało spadkiem lub też utratą aktywności enzymatycznej hAID (Rysunek 4.1.11. przedstawia badanie aktywności deaminazowej preparatu oczyszczonego na złożu glutationowym, a następnie poddanego oczyszczaniu metodą sączenia molekularnego, frakcje S1-S9 zawierały takie same ilości hAID, analizę frakcji S1-S8 przedstawiono na Rysunku 4.1.7). Do dalszych badań zdecydowano się zatem używać frakcji eluowanej z kolumny wypełnionej złożem glutationowym (frakcja E1, Rysunek 4.1.3.), w której uprzednio wymieniano bufor na bufor do badania aktywności hAID.



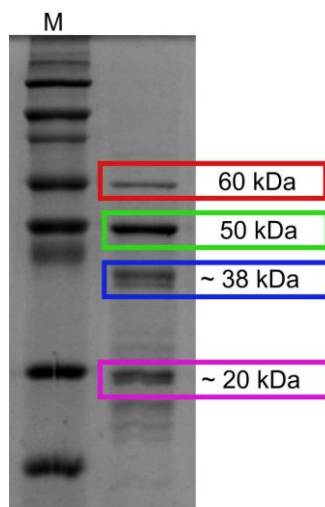
Rysunek 4.1.11. Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących produktów powstających w trójetapowym teście aktywności deaminazowej preparatu oczyszczonego na złożu glutationowym, a następnie poddanego oczyszczaniu metodą sączenia molekularnego. „K+” – pozytywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca U w pozycji 40); „K-” – negatywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca C w pozycji 40); E1 – trójetapowy test aktywności deaminazowej preparatu białkowego oczyszczonego tylko na złożu glutationowym; S0 – trójetapowy test aktywności deaminazowej preparatu białkowego przed sączeniem molekularnym; od S1 do S9 – trójetapowe testy aktywności deaminazowej frakcji uzyskanych w wyniku sączenia molekularnego.

4.2. Wstępna charakterystyka otrzymanego białka

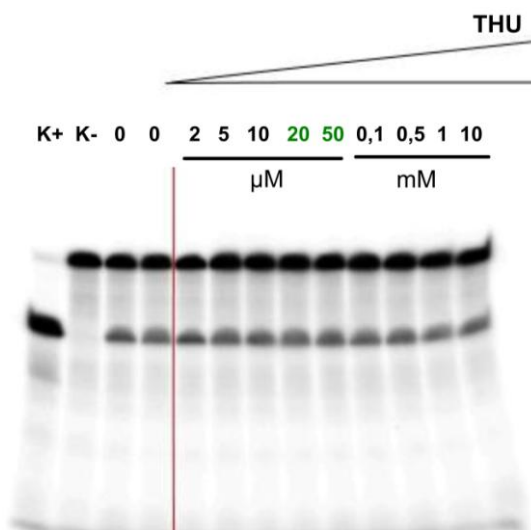
Dane literaturowe wskazują, że pierwsze preparaty AID otrzymywane w systemie prokariotycznym były zanieczyszczone deaminazami bakteryjnymi wykazującymi aktywność wobec wolnej cytydyny. W konsekwencji pierwsze dane na temat właściwości biochemicznych AID nie były spójne – część badaczy sądziła, że AID może deaminować wolną cytydynę i być inhibowana przez tetrahydrourydynę (THU). THU jest analogiem stanu przejściowego deaminacji wolnej cytydyny. Dziś większość badaczy jest zgodna co do tego, że zarówno AID jak i inne białka rodziny AID/APOBEC deaminują cytydynę jedynie w łańcuchach polinukleotydowych, a ich inhibicja przez THU jest wątpliwa [3].

W celu wykluczenia możliwości, iż obserwowana aktywność jest aktywnością enzymu bakteryjnego przeprowadzono szereg dodatkowych analiz. Aby potwierdzić, że uzyskane białko to hAID, przeprowadzono analizę metodą MALDI TOF. Jest to technika spektrometrii mas powszechnie stosowana w badaniach proteomicznych do identyfikacji białek. Spektrometry MALDI TOF bazują na technice jonizacji metodą desorpcji laserowej z udziałem matrycy (ang. *Matrix Assisted Laser Desorption Ionisation - MALDI*) oraz wykorzystują analizator czasu przelotu jonów (ang. *Time Of Flight, TOF*). Analizę przeprowadzono dla frakcji otrzymanej w wyniku oczyszczania na złożu glutationowym ekstraktu białkowego pochodzącego z bakterii BL21(DE3)pLysS transformowanych wektorem pGEX4T-1-hAID. Z żelu poliakrylamidowego, w którym analizowano frakcje po oczyszczaniu na złożu glutationowym wycinano fragment zawierający prążek odpowiadający najprawdopodobniej hAID (białko o masie 50 kDa) (Rysunek 4.2.1.). Z wyciętego fragmentu żelu eluowano białko i poddawano je analizie w Pracowni Metabolomiki i Proteomiki Instytutu Chemii Bioorganicznej PAN w Poznaniu. Badania potwierdziły, że wycięty fragment żelu zawiera białko hAID (wykryto peptydy pochodzące zarówno ze znacznika GST jak i z hAID). Dodatkowo poddano analizie 3 inne białka (o masie ok. 60, 38 i 20 kDa; Rysunek 4.2.1.) występujące w analizowanej frakcji w znacznej ilości. Wykazano, że białko o masie ok. 20 kDa jest fragmentem hAID, powstałym najprawdopodobniej wskutek degradacji. Białko o masie 60 kDa zidentyfikowano jako bakteryjne białko opiekuńcze (chaperonowe). Białka o masie ok. 38 kDa nie udało się zidentyfikować. Przeprowadzone analizy zmniejszyły prawdopodobieństwo, że

obserwowana aktywność deaminazowa preparatu może wynikać z obecności w preparacie deaminaz bakteryjnych.

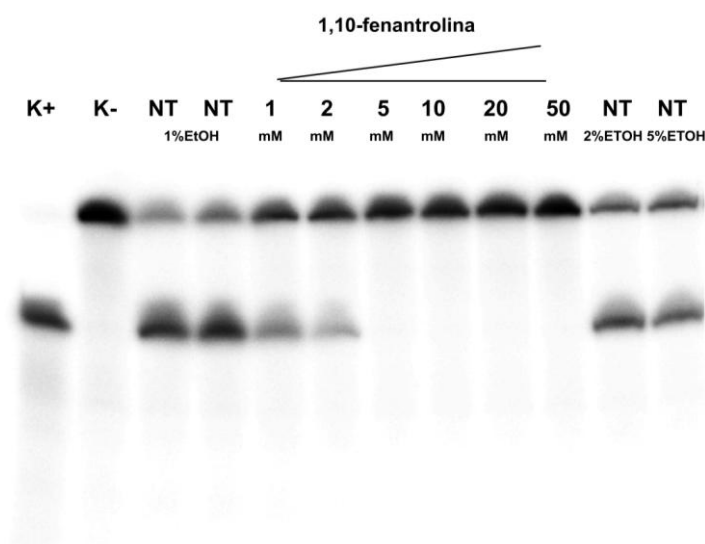


Rysunek 4.2.1. Analiza metodą elektroforezy w 12% żelu poliakrylamidowym w warunkach denaturujących frakcji otrzymanej w wyniku oczyszczania na złożu glutationowym ekstraktu białkowego pochodzącego z bakterii BL21(DE3)pLysS transformowanych wektorem pGEX4T-1-hAID. Fragmenty wycinane z żelu, zawierające uwidocznione w postaci prążków białka o masie ok. 50, 60, 38 i 20 kDa, zaznaczono kolorowymi prostokątami. Po elucji z fragmentów żelu, białka poddano analizie MALDI TOF. M – wzorzec mas.



Rysunek 4.2.2. Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących produktów powstających w trój etapowym teście aktywności deaminazowej prowadzonym w obecności THU w zakresie stężeń od 2 μM do 10 mM. „K+” – pozytywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca U w pozycji 40); „K-” – negatywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca C w pozycji 40); „0” – trój etapowy test aktywności deaminazowej prowadzony bez THU; kolejne ścieżki – trój etapowe testy aktywności deaminazowej prowadzone w obecności różnych stężeń THU (zielonym kolorem zaznaczono stężenia, które według danych literaturowych hamują działanie bakteryjnej deaminazy wolnej cytydyny).

W kolejnym etapie przeprowadzono testy inhibicji obserwowanej aktywności deaminazowej przez THU i 1,10-fenantrolinę. Jak już wspomniano THU jest analogiem stanu przejściowego reakcji enzymatycznych katalizowanych przez deaminazy wolnej cytydyny. 1,10-fenantrolina jest natomiast silnym czynnikiem kompleksującym metale przejściowe, co czyni ją silnym inhibitorem wszystkich deaminaz zależnych od cynku. Testy inhibicji przeprowadzono dla preparatu białkowego otrzymanego w wyniku oczyszczania na kolumnie wypełnionej złożem glutationowym w zakresie stężeń THU od 2 μM do 10 mM (dane literaturowe wskazują, że bakteryjne deaminazy wolnej cytydyny są inhibowane przez THU w stężeniu 20-50 μM [185]) oraz w zakresie stężeń 1,10-fenantroliny od 1 mM do 50 mM. Wyniki analiz przedstawiono na Rysunkach 4.2.2. oraz 4.2.3.

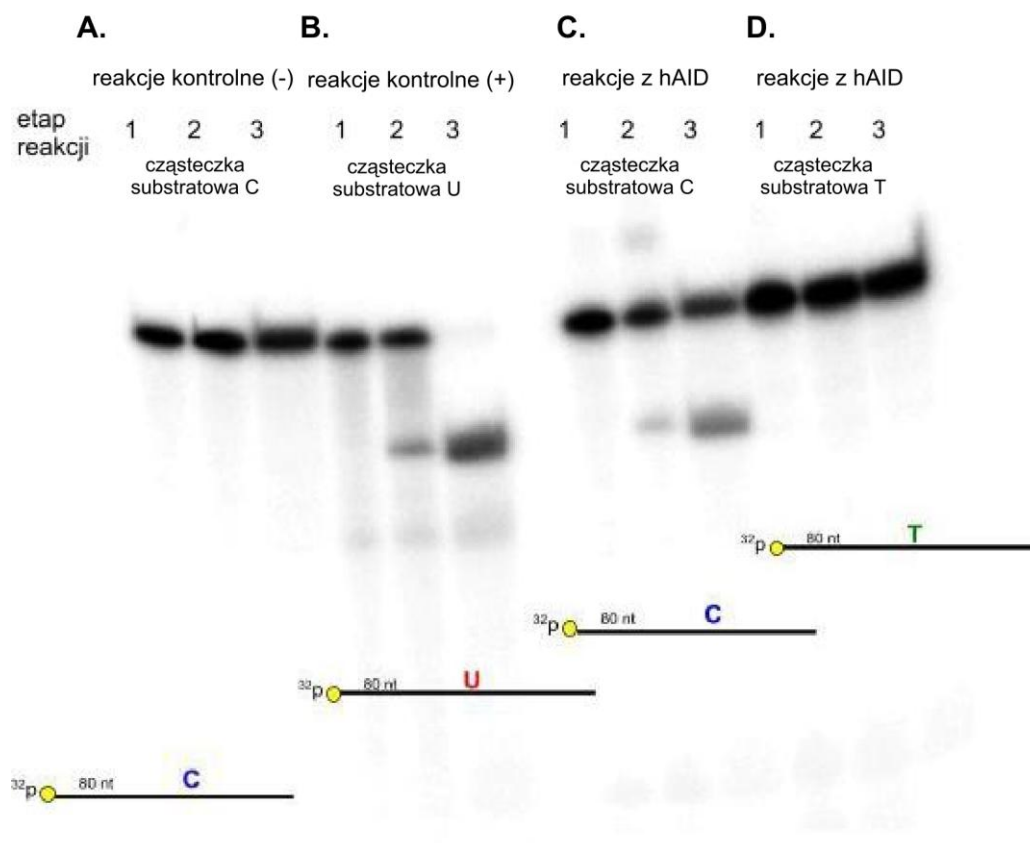


Rysunek 4.2.3. Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących produktów powstających w trój etapowym teście aktywności deaminazowej prowadzonym w obecności 1,10-fenantroliny w zakresie stężeń 1 – 50 mM. „K+” – pozytywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca U w pozycji 40); „K-” – negatywna reakcja kontrolna (bez preparatu białkowego, cząsteczka substratowa zawierająca C w pozycji 40); „NT 1%EtOH”, „NT 2%EtOH”, „NT 5%EtOH” – trój etapowy test aktywności deaminazowej prowadzony bez 1,10-fenantroliny, w obecności etanolu. Ponieważ 1,10-fenantrolina jest rozpuszczalna w etanolu, wykonano reakcje kontrolne, w których przetestowano wpływ etanolu na aktywność deaminazową badanego preparatu. W reakcjach tych obecny był etanol w stężeniu końcowym 1%, 2% lub 5% (stężenie 5% odpowiada stężeniu alkoholu w reakcji zawierającej 1,10-fenantrolinę w stężeniu 50 mM). Pozostałe ścieżki – trój etapowe testy aktywności deaminazowej prowadzone w obecności 1,10-fenantroliny, w stężeniach od 1 do 50 mM. Widoczna jest całkowita inhibicja aktywności deaminazowej hAID przez 1,10-fenantrolinę już w stężeniu 5 mM.

Stwierdzono, że obserwowana aktywność deaminazowa była całkowicie hamowana przez 1,10-fenantrolinę już w stężeniu 5 mM. Wynik ten sugeruje, że obserwowana aktywność była aktywnością deaminazy zależnej od cynku. Stwierdzono również, że obserwowana aktywność nie podlegała inhibicji przez THU. Świadczy to o tym, że obserwowana aktywność nie była aktywnością bakteryjnej deaminazy wolnej cytydyny.

Wybrany sposób badania aktywności hAID jest trzyetapową reakcją na drodze której z 80-nukleotydomowej cząsteczki ssDNA generowany jest produkt o długości 40 nukleotydów. Przyjęto uważać, że obecność 40-nukleotydomowego produktu świadczy o obecności aktywności deaminazowej. Jednak, w celu jednoznacznego potwierdzenia, że obserwowana aktywność jest aktywnością specyficzną wobec cytydyny, testowano aktywność preparatu białkowego wobec cząsteczki substratowej niezawierającej cytydyny (wykorzystana cząsteczka zawierała tymidynę w pozycji 40). Test ten wykazał brak wpływu badanego preparatu na cząsteczki DNA niezawierające cytydyny (Rysunek 4.2.4.), co świadczy o specyficzności reakcji. Testowano również produkty reakcji po poszczególnych etapach, w celu sprawdzenia po którym etapie pojawia się produkt końcowy o długości 40 nukleotydów. W przypadku zanieczyszczenia badanego preparatu nukleazami, można by się spodziewać obecności krótszych fragmentów DNA już po pierwszym etapie reakcji. Analiza wykazała obecność jedynie 80-nukleotydomowej cząsteczki substratowej po pierwszym etapie reakcji. Dowodzi to, że obserwowana aktywność nie jest aktywnością nukleazową. Pojawienie się niewielkiej ilości 40-nukleotydomowego produktu reakcji po drugim etapie można tłumaczyć podatnością miejsca apirymidynowego na rozerwanie nici DNA.

Opisane powyżej doświadczenia wskazują, że aktywność deaminazowa otrzymanego preparatu białkowego jest aktywnością hAID, i że charakterystyka biochemiczna tej aktywności jest spójna z najnowszymi danymi literaturowymi.



Rysunek 4.2.4. Ocena specyficzności działania otrzymanego preparatu hAID. Analiza metodą elektroforezy w 15% żelu poliakrylamidowym, w warunkach denaturujących produktów powstających w trój etapowym teście aktywności deaminazowej. Produkty analizowano po każdym z trzech etapów testu badania aktywności. A) reakcja kontrolna (-) - bez preparatu białkowego, cząsteczka substratowa zawierająca C w pozycji 40. B) reakcja kontrolna (+) - bez preparatu białkowego, cząsteczka substratowa zawierająca U w pozycji 40. C) reakcja z otrzymanym preparatem hAID, cząsteczka substratowa zawierająca C w pozycji 40. D) reakcja z preparatem hAID, cząsteczka substratowa zawierająca T w pozycji 40 (nie zawiera ona w swej sekwencji cytydyny).

4.3. Wybór i otrzymywanie mutantów hAID

W celu zbadania mechanizmów molekularnych leżących u podłoża aktywności biologicznej AID, otrzymano trzy mutanty hAID. Spośród wielu scharakteryzowanych mutantów AID (Tabela 1.4.1) wybrano trzy: R50A, N51A, Δ 189. W przypadku R50A arginina w pozycji 50 zastąpiona została przez alaninę, a w przypadku N51A asparagina w pozycji 51 zastąpiona została również przez alaninę. W obu mutantach pojedyncza substytucja aminokwasowa dotyczyła reszty ułożonej w pobliżu centrum katalicznego hAID. Mutant Δ 189 był wariantem białka skróconym o dziewięć reszt aminokwasowych z końca C. Wcześniejsze badania wskazywały, że wybrane mutanty wykazywały *in vitro* zróżnicowaną aktywność deaminazową wobec niemodyfikowanej cytydyny: R50A obniżoną, N51A całkowity jej brak a Δ 189 wyższą niż aktywność WT hAID. Co ciekawe, zebrane dane literaturowe (Tabela 4.3.1) wskazują, że aktywność *in vitro* tych mutantów nie koreluje z wydajnością co najmniej jednego z procesów inicjowanych przez AID (SHM lub CSR) w warunkach *in vivo*. W komórkach zawierających testowane mutacje w genie *AID* wydajność procesów CSR/SHM wynosiła: 25%/130% dla R50A, 2%/58% dla N51A oraz 75%/7% dla Δ 189. Obserwowany brak korelacji pomiędzy aktywnością *in vitro*, a wydajnością procesu SHM i CSR *in vivo* sugeruje, że istnieją jeszcze nieznanne czynniki determinujące działanie AID. Co ważne, aktywność deaminazowa wybranych mutantów AID wobec modyfikowanych substratów jak dotąd nie była badana.

Tabela 4.3.1. Charakterystyka wybranych mutantów AID. Wpływ poszczególnych mutacji na wydajność procesów CSR i SHM *in vivo* oraz aktywność deaminazową AID wobec niemodyfikowanej cytydyny *in vitro*.

wariant AID	intensywność procesu CSR lub SHM dane z systemów <i>in vivo</i> (% względem typu dzikiego)		aktywność <i>in vitro</i> (% względem typu dzikiego)	źródło
	SHM	CSR		
R50A	130%	25%	22%	[35]
N51A	58%	2%	0%	[35]
Δ 189	7%	75%	>100%	[97], [99]
WT	100%	100%	100%	—

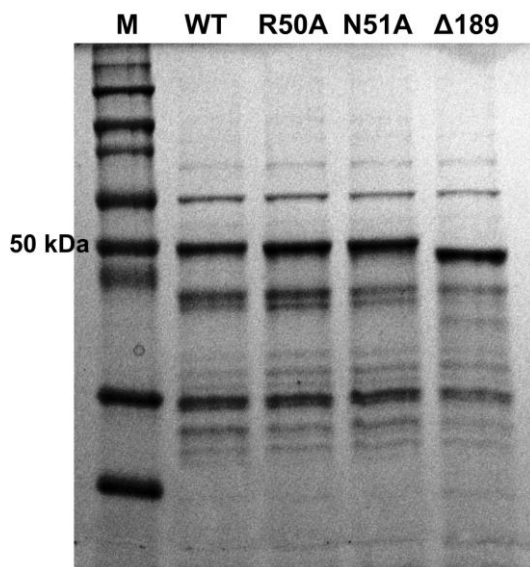
W identyczny sposób jak opisany powyżej dla WT hAID otrzymano trzy preparaty białkowe, zawierające mutanty R50A, N51A i Δ 189. Konstrukty kodujące

warianty hAID zawierające mutację punktową R50A lub N51A otrzymywano na drodze ukierunkowanej mutagenезy (podrozdział 3.2.3.2). Tą metodą wprowadzono jedną z powyższych, niesynonimicznych mutacji do otrzymanego wcześniej konstruktu pGEX4T-1-hAID. Metoda ukierunkowanej mutagenезy polega na amplifikacji całego plazmidu przy użyciu dwóch wzajemnie komplementarnych starterów wprowadzających odpowiednie mutacje. Były to startery mutR50A_F i mutR50A_R dla mutantu R50A (wprowadzające kodon GCC kodujący alaninę zamiast CGC kodującego argininę) oraz startery mutN51A_F i mutN51A_R dla mutantu N51A (wprowadzające kodon GCT kodujący alaninę zamiast AAT kodującego asparaginę) (Tabela 3.5.). Jako matrycę dla tej reakcji wykorzystywano wektor kodujący WT hAID (pGEX4T-1-hAID). Do amplifikacji plazmidu wykorzystywano polimerazę o wysokiej wierności kopiowania – polimerazę DNA Pfu. Produkty reakcji oczyszczano, a następnie poddawano trawieniu enzymem DpnI. Enzym ten trawi metylowany DNA. Pozwala on zatem na usunięcie z mieszaniny reakcyjnej matrycy użytej w reakcji PCR (produkty reakcji PCR pozostają niestrawione). Produkty reakcji trawienia ponownie oczyszczano, a uzyskanym DNA transformowano komórki kompetentne *E.coli* szczepu One shot[®] TOP10 (podrozdział 3.2.3.3.). Z uzyskanych kolonii bakteryjnych izolowano plazmidy w małej skali (podrozdział 3.2.1.10.), które następnie sekwencjonowano (podrozdział 3.2.1.6.). Zidentyfikowano plazmidy zawierające oczekiwaną mutację. Jeden z nich posłużył do transformacji szczepu bakterii BL21(DE3)pLysS stosowanego do produkcji białka.

Konstrukt kodujący wariant hAID skrócony o 10 kodonów na końcu 3' (mutant Δ 189) otrzymywano metodą PCR. Jako matrycę ponownie zastosowano plazmid pGEX4T-1-hAID. W reakcji użyto startery: F_04_aid oraz mut Δ 189_R (Tabela 3.5.). Powielony fragment genu hAID poddawano cięciu restrykcyjnemu i ligacji z wektorem ekspresyjnym pGEX4T-1. Procedura klonowania przebiegała identycznie jak w przypadku DNA kodującego WT hAID. Poprawność otrzymanego konstruktu potwierdzono poprzez sekwencjonowanie. Uzyskanym konstruktem pGEX4T-1- Δ 189 transformowano komórki kompetentne BL21(DE3)pLysS stosowane do produkcji białka.

Dalsze etapy przygotowania preparatów zawierających zmutowane warianty białka hAID przebiegały identycznie jak w przypadku WT hAID. Preparaty białkowe pochodzące z bakterii transformowanych jednym z przygotowanych konstruktyw oczyszczano na kolumnie wypełnionej złożem glutationowym. Eluowane z kolumny

frakcje analizowano metodą elektroforezy w żelu poliakrylamidowym w warunkach denaturujących (Rysunek 4.3.1.). Frakcje te poddawano wymianie buforu na bufor do badania aktywności hAID i wykorzystywano w dalszych badaniach.



Rysunek 4.3.1. Analiza w 12% żelu poliakrylamidowym, w warunkach denaturujących preparatów białkowych oczyszczonych na kolumnie wypełnionej złożem glutationowym. Każdy z preparatów zawierał jeden z testowanych wariantów hAID: WT, R50A, N51A, Δ189. Białka o masie ok. 50 kDa odpowiadają poszczególnym wariantom hAID. Zgodnie z oczekiwaniami, mutant Δ189 posiada nieznacznie mniejszą masą od pozostałych wariantów. M – wzorzec mas.

4.4. Charakterystyka aktywności hAID i jej mutantów w systemie *in vitro*

4.4.1. Charakterystyka aktywności deaminazowej hAID i jej mutantów wobec pojedynczych motywów: WRCY, WR5mCY, WR5hmCY¹⁶

Podstawowym substratem dla hAID jest cytydyna (C). Źródła literaturowe nie wskazują jednoznacznie czy enzym ten zachowuje się analogicznie wobec jej modyfikowanych pochodnych takich jak 5-metylocytydyna (5mC) i 5-hydroksymetylocytydyna (5hmC). W celu lepszego poznania uwarunkowań strukturalnych aktywności hAID, testowano jak zmienia się zdolność wszystkich czterech uzyskanych wariantów (WT, R50A, N51A oraz Δ 189) do deaminacji C, 5mC i 5hmC. W przeprowadzonych testach aktywności wykorzystano 80-nukleotydowe cząsteczki ssDNA (Tabela 3.5.) znakowane radioizotopowo na końcu 5' (podrozdział 3.2.4.1.). Cząsteczki te zawierały tylko jedną resztę C, 5mC lub 5hmC w pozycji 40, zlokalizowaną w obrębie czteronukleotydowego motywu AGCT, AG5mCT lub AG5hmCT. Motyw ten jest, według danych literaturowych, preferowany przez WT hAID [32]. Aby wykryć specyficzne dla każdego z tych substratów produkty deaminacji, stosowano jedną z trzech opracowanych wcześniej metod [138, 176, 177]. Metodę badania aktywności hAID wobec pojedynczej cytydyny znajdującej się w obrębie motywu AGCT opisano w podrozdziale 4.1 („Wstępne testy aktywności deaminazowej hAID”). W celu sprawdzenia czy produkowane białko wykazuje aktywność deaminazową wobec pojedynczego motywu AG5mCT stosowano zmodyfikowaną wersję tej metody. Cząsteczkę substratową zawierającą 5mC (Oligo_5mC, Tabela 3.5.) inkubowano 30 minut w temperaturze 37°C z testowanym preparatem białkowym (w ilości 4,5 μ g całkowitego białka). W przypadku obecności aktywności deaminazowej na tym etapie dochodzi do deaminacji 5mC do tymidyny (T). W kolejnym etapie do reakcji dodawano KCl do stężenia końcowego 50 mM oraz 30 pmola oligonukleotydu komplementarnego do badanej cząsteczki substratowej (Oligo_C-komplement, Tabela 3.5.). Mieszaninę reakcyjną inkubowano 5 min w temperaturze 95°C co skutkowało denaturacją oligonukleotydów. Następnie powoli (1°C / min) schładzano mieszaninę do temperatury 60°C. W sytuacji gdy na pierwszym etapie doszło do deaminacji, uzyskiwany był dsDNA z pojedynczym niesparowaniem G:T w pozycji 40. Do reakcji dodawano następnie glikozylazę TDG (Trevigene) i inkubowano dalsze 2 h w temperaturze 60°C. W przypadku obecności pojedynczego

¹⁶ W = A/T; R = A/G; Y = C/T

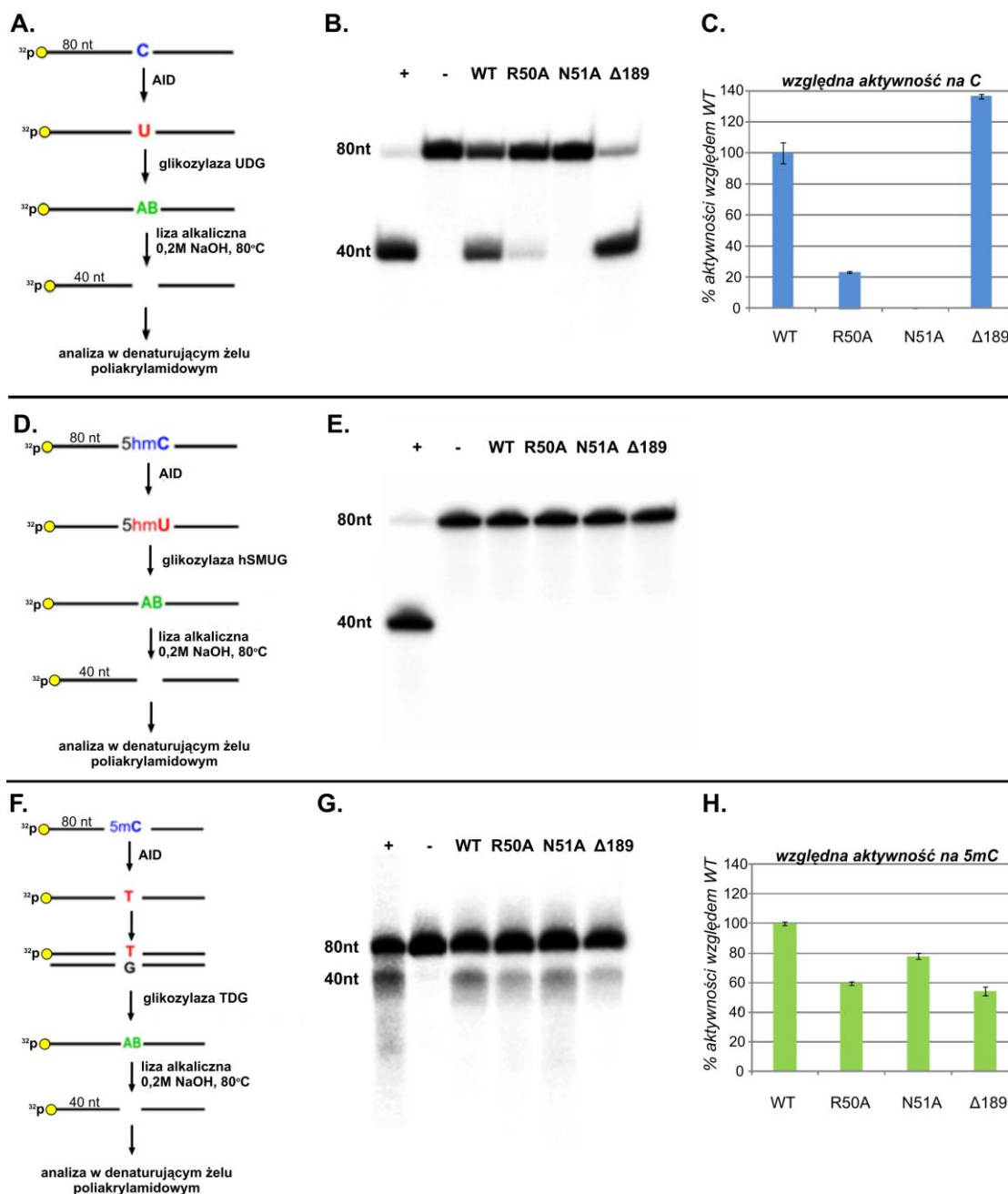
niesparowania G:T dochodzi do depirymidynacji T nietworzącej komplementarnej pary z G (enzym TDG usuwa tyminę, pozostawiając miejsce AP). Następnie przeprowadzano hydrolizę alkaliczną wiązania fosfodiesterowego w miejscu AP. Produkty reakcji analizowano metodą elektroforezy w 15% żelach poliakrylamidowych w warunkach denaturujących (podrozdział 3.2.1.4.). W przypadku obecności aktywności deaminazowej spodziewano się zaobserwować produkt o długości 40 nukleotydów. Równocześnie przeprowadzono reakcje kontrolne: (i) pozytywną reakcję kontrolną, bez preparatu białkowego, ale z wykorzystaniem cząsteczki substratowej posiadającej T zamiast C w pozycji 40 (Oligo_T, Tabela 3.5); (ii) negatywną reakcją kontrolną, bez preparatu białkowego, ale z wykorzystaniem standardowej cząsteczki substratowej (Oligo_5mC).

W celu sprawdzenia czy produkowane białko wykazuje aktywność deaminazową wobec pojedynczego motywu AG5hmCT stosowano analogiczną metodę do metody badania aktywności wobec motywu AGCT z tą jednak różnicą, iż w drugim etapie reakcji glikozylazę UDG zastąpiono glikozylazą hSMUG1 (New England Biolabs). Glikozylaza ta rozpoznaje i usuwa 5-hydroksymetylowouracyl pozostawiając miejsce AP. W metodzie tej stosowano również odpowiednio zmienioną cząsteczkę substratową (Oligo_5hmC, Tabela 3.5.), a w pozytywnej reakcji kontrolnej wykorzystywano cząsteczkę substratową posiadającą 5hmU zamiast 5hmC w pozycji 40 (Oligo_5hmU, Tabela 3.5). Porównanie wszystkich trzech metod badania aktywności wobec pojedynczych motywów AGCT, AG5mCT lub AG5hmCT przedstawia Rysunek 4.4.1.

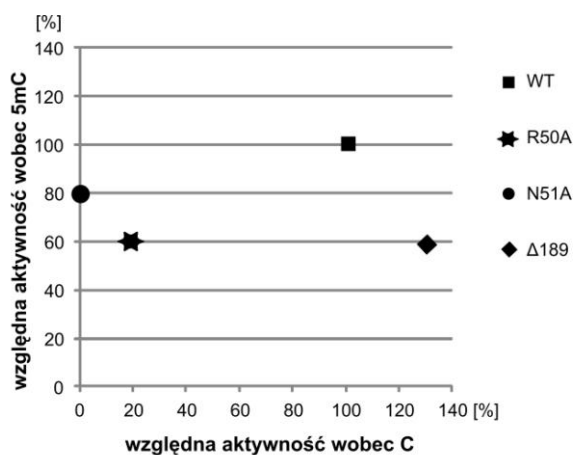
W wyniku przeprowadzonych analiz potwierdzono, że aktywność badanych wariantów hAID wobec niemodyfikowanej cytydyny jest zgodna z danymi literaturowymi. WT hAID wykazywała w testowanych warunkach wysoką aktywność deaminazową (przyjętą w obliczeniach jako 100%), jej mutant $\Delta 189$ wykazywał aktywność wyższą niż WT hAID (140% aktywności), mutant R50A wykazywał zaledwie 20% aktywność w porównaniu z WT hAID, a mutant N51A nie wykazywał aktywności deaminazowej (Rysunek 4.4.1.A – C). W kolejnym etapie badano aktywność deaminazową testowanych wariantów hAID wobec modyfikowanej cytydyny znajdującej się w obrębie tego samego kontekstu sekwencyjnego. Jak przedstawiono na Rysunku 4.4.1.D – E żaden z badanych wariantów hAID nie wykazywał aktywności wobec 5hmC. Co ciekawe, wszystkie badane białka wykazywały aktywność wobec metylowanego substratu (Rysunek 4.4.1.F – H).

Analogicznie jak uprzednio przyjęto, że aktywność WT hAID wynosi 100% i w odniesieniu do niej wyznaczono aktywność mutantów. Stwierdzono, że aktywność deaminazowa testowanych wariantów hAID wobec 5mC nie korelowała jednak z ich aktywnością wobec C (Rysunek 4.4.2.). Obliczony współczynnik korelacji liniowej wynosił 0,066¹⁷. Nadaktywny wobec C mutant $\Delta 189$ wykazywał wobec 5mC jedynie 50% aktywności. Mutant R50A wykazywał obniżoną aktywność wobec 5mC. Obserwowany spadek aktywności był jednak mniej wyraźny niż w przypadku C. Nieoczekiwanie okazało się, że mutant N51A nieaktywny wobec C wykazywał znaczącą aktywność wobec 5mC.

¹⁷ Wartość współczynnika korelacji liniowej mieści się w przedziale od -1 do 1. Wartość -1 odpowiada maksymalnej korelacji ujemnej. Wartość 1 odpowiada maksymalnej korelacji dodatniej. Wartość 0 oznacza brak korelacji.



Rysunek 4.4.1. Badanie aktywności deaminazowej WT hAID oraz jej mutantów R50A, N51A, Δ189 wobec C (A-C), 5hmC (D-E) oraz 5mC (F-H). W podpunktach A, D, F przedstawiono metody badania aktywności deaminazowej wobec poszczególnych substratów. W podpunktach B, E, G przedstawiono analizy w denaturujących żelach poliakrylamidowych produktów powstających w testach aktywności deaminazowej wobec poszczególnych substratów przez badane warianty hAID: WT, R50A, N51A, Δ189. W przypadku obecności aktywności deaminazowej, widoczny jest produkt o długości 40 nt. „+” – pozytywna reakcja kontrolna; „-”, – negatywna reakcja kontrolna (opis w tekście). W podpunktach C i H przedstawiono wykresy aktywności badanych mutantów hAID wobec C lub 5mC. Aktywności wyrażono w % względem aktywności WT hAID. Aktywności poszczególnych wariantów hAID szacowano na podstawie stosunku intensywności prążka 40 nt do prążka 80 nt w obrębie każdej ze ścieżek. Analizy tej dokonano przy pomocy programu MultiGauge V3.0 na podstawie wyników uzyskanych z trzech niezależnych doświadczeń.



Rysunek 4.4.2. Wykres przedstawiający brak korelacji pomiędzy względną aktywnością deaminazową badanych mutantów hAID wobec C i 5mC. Przyjęto, że aktywność WT hAID wynosi 100%. Aktywność deaminazowa każdego z mutantów wyznaczona została w procentach względem aktywności WT hAID.

4.4.2. Charakterystyka aktywności deaminazowej hAID i jej mutantów

N51A wobec matryc zawierających wszystkie sekwencje typu NNCN

Brak aktywności deaminazowej mutantów N51A wobec motywu WRCY można przypisać: (i) niezdolności tego mutantów do deaminacji niemetylowanej cytydyny lub (ii) zmianie jego preferencji wobec motywu sekwencyjnego w ramach którego cytydyna ulega deaminacji. Aby stwierdzić, która z powyższych możliwości jest prawdziwa postanowiono zbadać aktywność deaminazową WT hAID i mutantów N51A wobec wszystkich możliwych motywów NNCN (gdzie N oznacza dowolny nukleotyd: A, G, C lub T) (podrozdział 3.2.4.5.). W tym celu zsyntetyzowano na drodze chemicznej cztery oligonukleotydy (każdy o długości 104 nukleotydów), zawierające wszystkie możliwe kombinacje motywów NNCN (4^4 czyli 64 kombinacje), oskrzydłone 20-nukleotydowymi odcinkami rozpoznawanymi przez startery (patrz oligonukleotydy NNCN_1, NNCN_2, NNCN_3 lub NNCN_4, Tabela 3.5.). Każdy z oligonukleotydów zawierał 16 kombinacji motywu NNCN. 0,3 pmola oligonukleotydu rozpuszczono w jednokrotnie stężonym buforze do badania aktywności hAID, poddano denaturacji, a następnie inkubowano przez 30 minut w temperaturze 37°C z testowanym preparatem białkowym (w ilości 4,5 µg całkowitego białka). W kolejnym etapie 1 µl mieszaniny reakcyjnej używano jako matrycy w reakcji PCR prowadzonej z wykorzystaniem gotowego zestawu odczynników KAPA HiFi HotStart Uracil+ ReadyMix (KAPA Biosystems). Zastosowany w trakcie amplifikacji enzym jest polimerazą o wysokiej

wierności kopiowania, która toleruje obecność uracylu w amplifikowanej cząsteczce DNA. Produkty PCR poddawano reakcji dodawania niesparowanych reszt adenozytowych do końców 3', a następnie ligowano z wektorem pGEM-T Easy (podrozdział 3.2.1.9.). Produktem ligacji transformowano komórki kompetentne *E. coli* szczep One shot[®] TOP10 (podrozdział 3.2.3.3.). Dwadzieścia niezależnych losowo wybranych kolonii bakteryjnych zawierających plazmid z insertem (których selekcji dokonano na podstawie testu α -komplementacji) wykorzystano do izolacji plazmidów w małej skali (podrozdział 3.2.1.10.), a następnie wyizolowane plazmidy sekwencjonowano (podrozdział 3.2.1.6.). Wyniki sekwencjonowania poddawano analizie bioinformatycznej, w której dokonano przyrównania sekwencji otrzymanych klonów do sekwencji odpowiedniego wyjściowego oligonukleotydu (NNCN_1 – NNCN_4). Obecność tranzycji cytydyny do tymidyny świadczyła o zajściu reakcji deminacji. Wyniki analiz bioinformatycznych przedstawiono na Rysunku 4.4.2.1. oraz w Załączniku 1. Dla WT hAID obserwowano zarówno cząsteczki ssDNA, w których wszystkie lub większość cytydyn uległa deaminacji jak i takie, które pozostały prawie niezmienione. Dla mutantu N51A obserwowano jedynie cząsteczki ssDNA o niezmienionej sekwencji.

Na podstawie przeprowadzonej analizy stwierdzono, że WT hAID ma zdolność do deaminacji wszystkich reszt C w ramach pojedynczej cząsteczki DNA (w obrębie różnych kontekstów sekwencyjnych) przed oddysocjowaniem i zadziałaniem na inną cząsteczkę DNA. Obserwacja ta potwierdza wcześniejsze doniesienia literaturowe grupy badawczej profesora Goodmana, w których postulowano, że enzym ten działa procesywnie [83, 186, 187]. W przeciwieństwie do WT hAID, mutant N51A nie wykazywał aktywności deaminazowej wobec żadnego z 64-możliwych motywów NNCN. Obserwacja ta dowodzi, że mutant N51A cechuje się zupełnym brakiem aktywności deaminazowej wobec niemetylowanej cytydyny. Na tej podstawie można wnioskować, iż niewielka zmiana struktury hAID w bliskim sąsiedztwie centrum katalicznego tego enzymu całkowicie znosi jego aktywność deaminazową wobec niemetylowanego substratu, zachowując jednocześnie jego aktywność wobec substratu metylowanego.

Aktywność WT hAID wobec czterech testowanych nukleotydów różniła się znacząco. Najwyższą aktywność stwierdzono dla oligonukleotydów NNCN_1 i NNCN_3, a najniższą dla oligonukleotydu NNCN_4. Otrzymane wyniki sugerują, że obserwowane różnice w aktywności wynikają raczej ze struktury drugorzędowej

cząsteczki substratowej (i dostępności poszczególnych cytydyn dla enzymu) niż z kontekstu sekwencyjnego w jakim znajduje się deaminowana cytydyna.

A.

NNCN_1 AGCAGACGACCCACGA-GCTGACCTTCACACCAGCCGACTTTCCCACTAGCGGACATTCCACA

klon 1 .. T ~
klon 2 .. T.. T.. T.. T.. ~. T.. T.. T.. TT. T.. TTT. T.. T.. T.. T..
klon 3 .. T.. TTT. T.. ~. T.. T.. T.. T.. T.. T.. TTT. T.. T.. T.. T..
klon 4 .. T.. T.. TTTT. T.. ~. T.. TT.. T.. T.. TT.. TT.. T.. TTT. T.. T.. T.. T..
klon 5 .. T.. T.. TTTT. T.. ~. T.. TT.. T.. T.. TT.. TT.. T.. TTT. T.. T.. T.. T..
klon 6 .. T.. T.. TTTT. T.. ~. T.. TT.. T.. T.. TT.. TT.. T.. TTT. T.. T.. T.. T..
klon 7 .. T.. TTT. T.. ~. T.. TT.. T.. T.. TT.. TT.. T.. TTT. T.. T.. T.. T..
klon 8 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 9 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 10 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 11 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 12 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 13 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 14 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 15 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 16 .. T.. T.. TT.. T.. A T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 17 .. T.. T.. TT.. T.. T.. ~. T.. T.. T.. T.. TT.. T.. T.. TTT. T.. T.. A T.. T.. T..
klon 18 .. T.. T.. TTTT. T.. ~. T.. TT.. T.. T.. TT.. T.. T.. T.. T.. T.. T.. T..
klon 19 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..
klon 20 .. T.. T.. T.. T.. ~. T.. T.. T.. T.. T.. T.. T.. T.. T.. T.. T..

B.

NNCN_1 AGCAGACG-ACCCACGAGCTGACC-TTCACACCAGCCG-ACTTTCCCACTAGCGGACATTCCACA

klon 1 .. T ~
klon 2 .. T ~
klon 3 .. T ~
klon 4 .. T ~
klon 5 .. T ~
klon 6 .. T ~
klon 7 .. T ~
klon 8 .. T ~
klon 9 .. T ~
klon 10 .. T ~
klon 11 .. T ~
klon 12 .. T ~
klon 13 .. T ~
klon 14 .. T.. G.. T.. C.. T.. G..
klon 15 .. T ~
klon 16 .. T ~
klon 17 .. T ~
klon 18 .. T ~
klon 19 .. T ~
klon 20 .. T ~

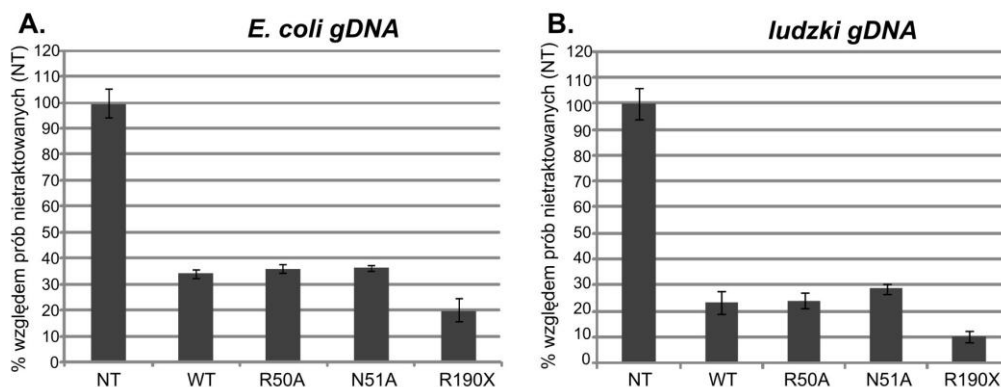
Rysunek 4.4.2.1. Porównanie sekwencji dwudziestu klonów oligonukleotydu NNCN_1 poddanego działaniu WT hAID (A) lub jej mutantu N51A (B). Znakiem „~” zaznaczono delecje / insercje wprowadzone przypuszczalnie przez zastosowaną podczas klonowania polimerazę; „.” oznaczono pozycje tożsame z sekwencją wyjściową; w miejscu zmiany widoczne są jednoliterowe symbole odpowiednie dla danego nukleotydu. Na rysunku nie przedstawiono sekwencji oskrzydających rozpoznawanych przez startery stosowane w reakcji PCR (zmiany w ich obrębie nie podlegają detekcji).

Testowane oligonukleotydy cechowały się podobnymi temperaturami topnienia. Wynosiły one: 93,6°C; 93,4°C; 95,7°C; 95,7°C odpowiednio dla oligonukleotydów od NNCN_1 do NNCN_4. Podjęto próbę bioinformatycznej analizy struktury drugorzędowej badanych cząsteczek ssDNA (zastosowano parametr temperaturowy – 37°C ponieważ w takiej temperaturze prowadzona była reakcja deaminacji). Mnogość wygenerowanych struktur dla tak długich cząsteczek ssDNA uniemożliwiła jednak znalezienie jakiegokolwiek korelacji pomiędzy strukturą drugorzędową cząsteczki substratowej, a aktywnością WT hAID. Konieczne jest zatem przeprowadzenie dalszych badań eksperymentalnych, aby wytłumaczyć obserwowane zjawisko. Wykracza to jednak poza przedmiot niniejszej pracy.

4.4.3. Badanie wpływu hAID i jej mutantów na poziom 5mC w genomowym DNA *in vitro*

W kolejnym etapie postanowiono zbadać wpływ otrzymanych wariantów hAID na globalny poziom 5mC w genomowym DNA. Doświadczenie to miało na celu określenie aktywności deaminazowej poszczególnych białek wobec 5mC w kontekście całego genomu, czyli wobec długich odcinków ssDNA zawierających liczne 5mC. Zmiany w globalnym poziomie 5mC w genomowym DNA (gDNA) w warunkach *in vitro* analizowano z wykorzystaniem gotowego zestawu odczynników 5mC DNA ELISA Kit (Zymo Research) zgodnie z instrukcją producenta (podrozdział 3.2.4.6). Metoda ta pozwala na względną ocenę zmian zawartości 5mC w badanej próbce DNA dzięki zastosowaniu monoklonalnych przeciwciał specyficznych wobec 5mC. Detekcję przeprowadzono metodą kolorymetryczną – przeciwciała II-rzędowe były sprzężone z peroksydazą chrzanową. Jako substrat w reakcji wykorzystywano metylowany gDNA pochodzący z *E. coli* (wchodzący w skład w/w gotowego zestawu odczynników) lub metylowany, ludzki gDNA (dostępny komercyjnie, Zymo Research). Zarówno bakteryjny jak i ludzki gDNA to cząsteczki metylowane enzymatycznie *in vitro*, powszechnie stosowane jako próby kontrolne w badaniach metylacji. 100 ng gDNA denaturowano, a następnie inkubowano w temperaturze 37°C z preparatem białkowym (w ilości 4,5 µg całkowitego białka) zawierającym jeden z badanych wariantów hAID. Dla gDNA pochodzącego z *E.coli* stosowano inkubację przez 8 h. Dla ludzkiego gDNA stosowano inkubację przez 16 h (dłuższa inkubacja wynika z mniejszej gęstości wysp CpG w stosunku do wielkości genomu dla człowieka w porównaniu z organizmem

bakteryjnym). Jako kontrole negatywne stosowano próbki gDNA inkubowane z buforem reakcyjnym niezawierającym testowanego preparatu białkowego. Równocześnie identycznej procedurze poddawano próby ślepe (niezawierające gDNA). Wykonywano trzy niezależne powtórzenia techniczne dla każdej z prób. Kolejne etapy reakcji obejmujące: denaturację badanych prób, opłaszczanie płytek wielodołkowych, inkubację w buforze blokującym, inkubację z przeciwciałami i detekcję kolorymetryczną, przeprowadzano zgodnie z instrukcją producenta. Pomiaru absorpcji światła o długości fali 450 nm dokonywano za pomocą czytnika płytek VictorTM X4 firmy Perkin Elmer. Wartości absorpcji normalizowano wobec ślepych prób. Następnie dla trzech powtórzeń wyznaczano średnią arytmetyczną i odchylenie standardowe. Zmiany w globalnym poziomie 5mC w gDNA prób badanych wyrażono procentowo względem kontroli negatywnych. Na podstawie wyników analiz (Rysunek 4.4.3.1.) stwierdzono, że WT hAID i jej mutanty: R50A, N51A, Δ 189 w warunkach *in vitro* drastycznie redukują globalny poziom 5mC w gDNA zarówno ludzkim, jak i pochodzącym z *E. coli*. Globalny poziom 5mC w gDNA pochodzącym z *E. coli* inkubowanym z WT hAID, R50A lub N51A został zmniejszony o 70%. Globalny poziom 5mC w ludzkim gDNA inkubowanym z tymi wariantami uległ zmniejszeniu o 80%. Redukcja ta była jeszcze wyraźniejsza w przypadku mutanta Δ 189 i wynosiła odpowiednio 80% dla gDNA z *E. coli* lub 90% w przypadku ludzkiego gDNA. Na tej podstawie stwierdzono, że hAID jest zdolna do redukcji globalnego poziomu 5mC w gDNA *in vitro* bez udziału żadnych kofaktorów. Odkrycie to pokazuje również, że hAID może wydajnie deaminować 5mC w długich cząsteczkach ssDNA zawierających liczne 5mC.



Rysunek 4.4.3.1. Wpływ WT hAID oraz jej mutantów (R50A, N51A, Δ 189) na globalny poziom 5mC w gDNA pochodzącym z *E.coli* (A) oraz w ludzkim gDNA (B). Przyjęto, że poziom 5mC w nietraktowanym hAID preparacie gDNA wynosi 100% (NT). Względem próby nietraktowanej hAID wyznaczono procentowy poziom 5mC w preparatach traktowanych WT hAID lub jednym z jej mutantów.

4.5. Bioinformatyczna analiza struktury hAID oraz jej kompleksów z substratami

Brak korelacji między aktywnością badanych wariantów hAID wobec C i 5mC sugerował, że istnieją różne elementy strukturalne warunkujące rozpoznawanie i/lub wiązanie C i 5mC przez centrum katalityczne hAID. Zjawisko to było szczególnie ewidentne w przypadku mutantu N51A, który nie wykazywał aktywności wobec C ale był aktywny wobec 5mC.

Co ciekawe, dane literaturowe sugerują, że głównym czynnikiem odpowiedzialnym za niższą efektywności deaminacji 5mC w porównaniu z C przez WT AID jest istnienie zawady przestrzennej w kieszeni katalitycznej enzymu wobec podstawnika przy węglu 5 pierścienia cytozyny [123, 124, 138]. W kolejnym etapie badań postanowiono zatem sprawdzić tą hipotezę i zbadać różnice w oddziaływaniach WT hAID oraz mutantu N51A z modyfikowanym i niemodyfikowanym substratem.

Otrzymanie modeli struktur 3D WT hAID oraz jej mutantu N51A. Ponieważ struktura trzeciorzędowa (3D) AID jak dotąd nie została poznana, model tej struktury wygenerowano *in silico* na drodze modelowania homologicznego. Metoda ta opiera się na założeniu, że białka homologiczne (pochodzące od wspólnego przodka) zazwyczaj zachowują podobną strukturę pomimo obserwowanych różnic sekwencji

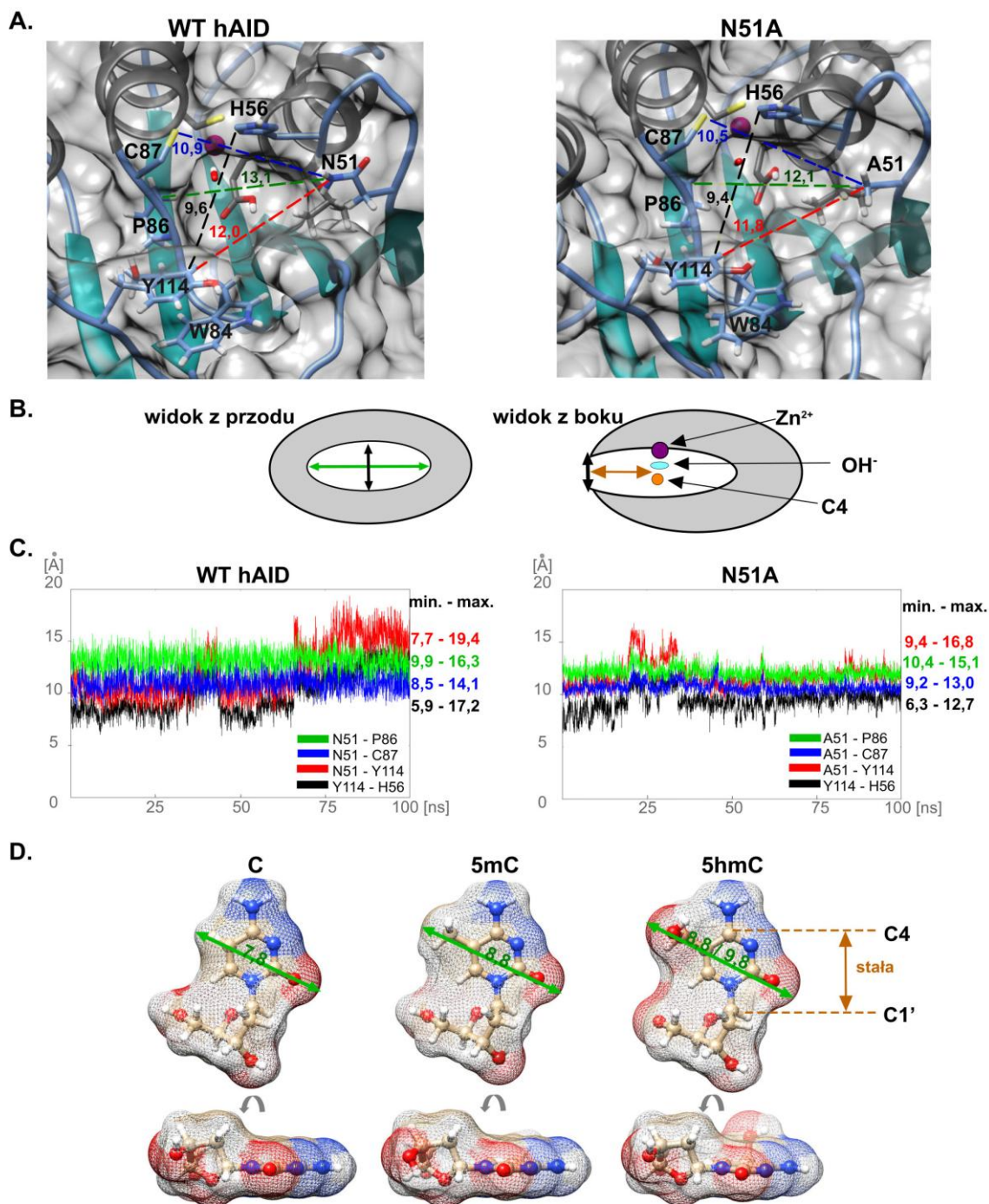
w aminokwasowej. Przyrównanie sekwencji badanego białka o nieznannej strukturze (celu) do sekwencji białka spokrewnionego o znanej strukturze (matrycy) jest punktem wyjścia do przeprowadzenia symulacji procesu ewolucji białka – celu z homologicznego białka – matrycy. Metoda ta jest obecnie najbardziej skuteczną strategią przewidywania struktur białek. Białkiem homologicznym – matrycą zastosowaną w procedurze modelowania struktury hAID była domena katalityczna białka APOBEC3G (dokładnie dwie struktury krystaliczne tej domeny zdeponowane w bazie PDB pod numerami: 3e1u i 3v4k). Modele struktury 3D WT hAID oraz mutantu N51A otrzymano dzięki uprzejmości profesora Janusza Bujnickiego z Międzynarodowego Instytutu Biologii Molekularnej i Komórkowej w Warszawie (podrozdział 3.2.5).

Uzyskane modele równoważono i badano ich stabilność metodą dynamiki molekularnej (MD). MD jest procesem wyznaczania trajektorii (ewolucji w czasie) układu cząsteczek na drodze całkowania równań ruchu Newtona, przy zadanych oddziaływaniach cząsteczkowych oraz odpowiednich warunkach początkowych. Umożliwia ona szczegółową analizę ruchów molekularnych i zmian konformacyjnych w obrębie badanego układu. Dynamika molekularna wygenerowanych modeli białek została przeprowadzona w Zakładzie Chemii i Biologii Strukturalnej Kwasów Nukleinowych ICHB PAN przez dr Joannę Sarzyńską. Wszystkie symulacje wykonano przy użyciu oprogramowania NAMD w polu siłowym *Charmm36* metodą dynamiki Langevin [182, 183]. Analiza 100 ns trajektorii MD obu modeli wykazała, że uzyskano ustabilizowane struktury bez znaczących zmian konformacyjnych.

Testowanie hipotezy o istnieniu zawady przestrzennej w centrum katalitycznym hAID wobec podstawnika przy węglu C5. W celu zbadania czy w centrach katalitycznych WT hAID oraz mutantu N51A istnieje zawada przestrzenna wobec podstawnika przy węglu C5 modyfikowanej cytydyny, dokonano szczegółowego pomiaru kieszeni katalitycznych WT hAID oraz mutantu N51A, z zastosowaniem oprogramowania Chimera 1.10.1. Jednocześnie określono wielkości modyfikowanych i niemodyfikowanych cytydyn. Wielkość kieszeni katalitycznych została zdefiniowana przez cztery odległości (Rysunek 4.5.1.A) między wybranymi atomami ciężkimi, tworzącymi wewnętrzną powierzchnię kieszeni katalitycznych. Atomy te należały do następujących par reszt aminokwasowych: (1) N51 (A51 w przypadku mutantu) i Y114; (2) N51 (A51) i P86; (3) N51 (A51) i C87; (4) Y114 i H56. Pierwsze trzy wybrane

odległości uwzględniały resztę N51, której substytucja alaniną skutkowała zmianą właściwości enzymatycznych białka. Czwarta odległość definiowała wysokość kieszeni katalitycznej. Wartość każdej z wyznaczonych odległości wyrażono jako średnią arytmetyczną wartości zmierzonych dla wszystkich klatek z trajektorii MD (Rysunek 4.5.1.A). Na Rysunku 4.5.1.C przedstawiono również zmiany tych wartości w obrębie trajektorii. Na podstawie powyższych analiz stwierdzono, że pomimo zasadniczych różnic we właściwościach enzymatycznych WT hAID i jej mutantu N51A, wielkości obu kieszeni katalitycznych były niemal identyczne. Zakres zmian wielkości wybranych wymiarów był większy w przypadku WT hAID. Centrum katalityczne N51A cechowała zatem nieco większa stabilność.

Wymiary reszt nukleotydowych determinujące możliwość wejścia danego substratu do kieszeni katalitycznej wynosiły odpowiednio: 7,8 Å dla C; 8,8 Å dla 5mC; 8,8 lub 9,8 Å dla 5hmC w zależności od rotacji grupy hydroksylowej (uwzględniając promienie van der Waalsa; Rysunek 4.5.1.C). Na podstawie powyższych pomiarów stwierdzono, że zarówno C jak i 5mC są mniejsze niż minimalne wartości wymiaru N51 (A51) – P86, zarówno w przypadku kieszeni katalitycznej WT hAID jak i N51A. Wyniki te sugerują zatem, że nie ma zawady przestrzennej w centrum katalitycznym WT hAID i jej mutantu N51A wobec podstawnika metylowego przy węglu C5.



Rysunek 4.5.1. Porównanie wielkości kieszeni katalitycznych WT hAID i mutantu N51A oraz reszt nukleotydowych C, 5mC i 5hmC. **A)** Wielkości kieszeni katalitycznych badanych wariantów hAID zostały zdefiniowane poprzez cztery odległości pomiędzy atomami następujących reszt aminokwasowych: N51 (A51 w przypadku mutantu) i Y114 (odległość zaznaczona czerwonym kolorem); N51 (A51) i P86 (kolor zielony); (3) N51 (A51) i C87 (kolor niebieski); (4) Y114 i H56 (kolor czarny). Przy każdej odległości podano jej wartość średnią z trajektorii MD wyrażoną w Å. Jako punkty pomiaru wybrano następujące atomy poszczególnych reszt aminokwasowych: ND2 atom reszty N51; CB atom reszty A51; CE2 atom reszty Y114; CB atom reszty P86 oraz SG atom reszty C87. **B)** Schemat kieszeni katalitycznej hAID – widok z przodu i widok z boku. Odległość pomiędzy N51 (A51) a P86 (odległość determinująca szerokość pierścienia pirymidynowego, który może wejść do kieszeni katalitycznej)

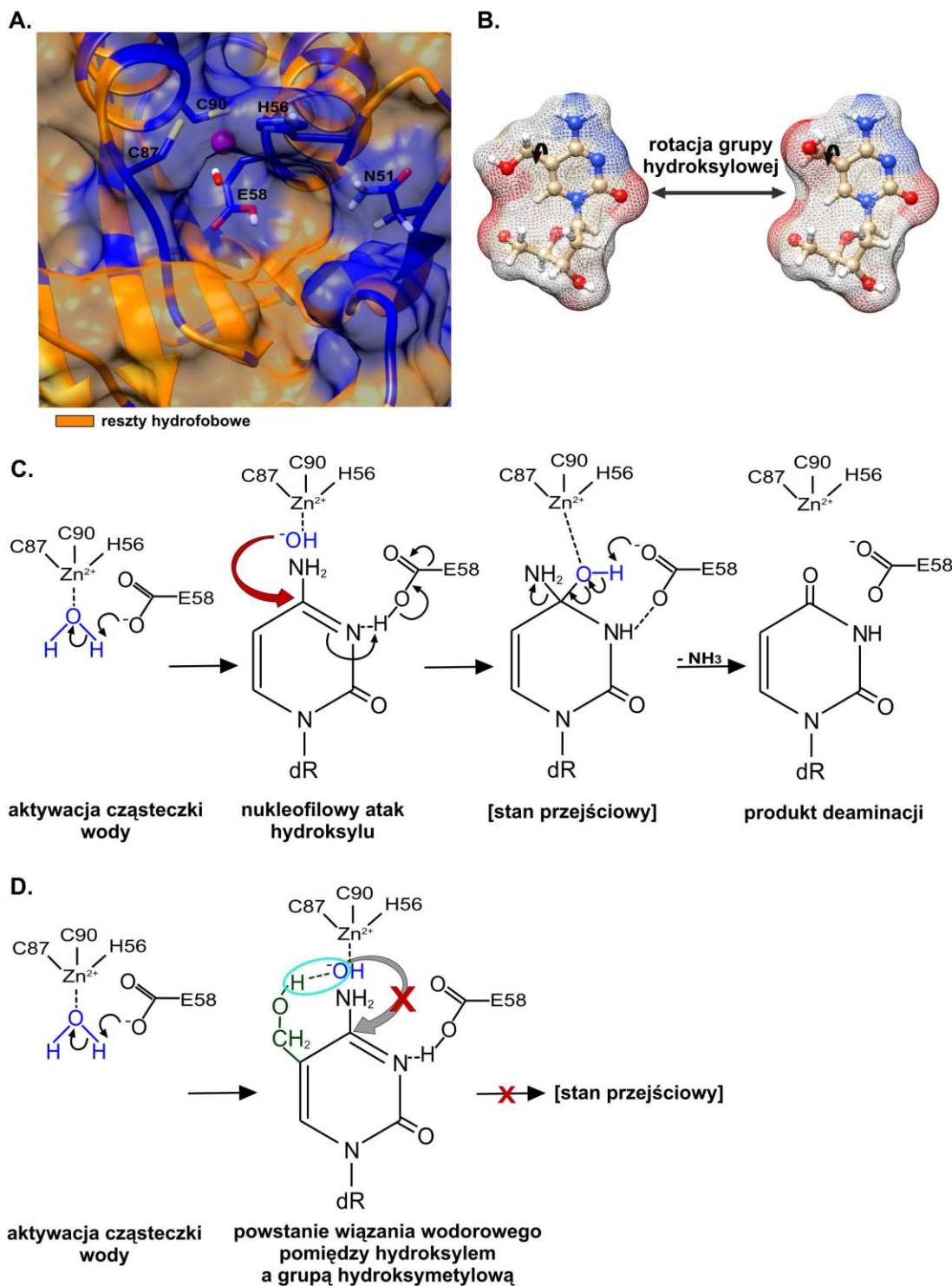
zaznaczono kolorem zielonym. Odległość determinującą wysokość kieszeni katalitycznej (odległość pomiędzy Y114 a H56) zaznaczono kolorem czarnym. Głębokość kieszeni katalitycznej zaznaczono kolorem pomarańczowym. **C)** Wykresy ilustrujące zmiany każdej ze zdefiniowanych odległości w obrębie trajektorii MD. Obok wykresów podano wartości minimalne i maksymalne z trajektorii dla każdej odległości. **D)** Wielkości cytydyny (C), 5-metylocytydyny (5mC) i 5-hydroksymetylocytydyny (5hmC) podane w Å dla wymiarów determinujących możliwość wejścia danego substratu do kieszeni katalitycznej (zielone strzałki). Dystans pomiędzy węglem C1 a C4 pierścienia pirymidynowego (pomarańczowa strzałka) jest wartością stałą i odpowiada głębokości kieszeni katalitycznych. Reszty nukleotydowe pokazano również w rzucie bocznym.

Wymiary 5hmC wskazują, że ona też powinna zmieścić się w kieszeni katalitycznej obu badanych enzymów (tylko w niewielkiej części trajektorii MD wymiary tych kieszeni były zbyt małe). Szukano zatem wyjaśnienia dla obserwowanego we wcześniejszych eksperymentach *in vitro* braku aktywności testowanych wariantów hAID wobec 5hmC. Na podstawie szczegółowych analiz uzyskanych modeli białek i danych literaturowych zaproponowano kilka możliwych przyczyn. Po pierwsze należy brać pod uwagę, że silnie hydrofilowy charakter grupy hydroksylowej sprawia, iż ulega ona hydratacji w roztworach wodnych. Grupa hydroksylowa przy węglu C5 może tworzyć wiązania wodorowe nawet z pięcioma cząsteczkami wody, a pojedyncza cząsteczka wody zwiększa wymiar hydratowanej grupy o 2,8 do 3,2 Å [188-191]. Rzeczywiste wymiary 5hmC są zatem wystarczająco duże, aby wykluczyć możliwość jej umiejscowienia w centrum aktywnym hAID. Dodatkowo, rotacja grupy hydroksylowej sprawia, że pierścień 5hmC, w przeciwieństwie do pierścienia 5mC i C, nie jest płaski (Rysunek 4.5.1. C), co jest dodatkowym czynnikiem utrudniającym prawidłowe ułożenie w kieszeni katalitycznej. Po drugie, dehydratacja niezbędna do zajścia deaminacji byłaby zjawiskiem energetycznie niekorzystnym, ze względu na konieczność rozerwania wiązań wodorowych pomiędzy grupą hydroksylową, a cząsteczkami wody. Po trzecie, jeśli nawet nastąpiłaby dehydratacja, to wówczas hydrofobowe reszty aminokwasowe tworzące „podstawę” kieszeni katalitycznej (Rysunek 4.5.2) wymusiłyby rotację grupy hydroksylowej w kierunku aktywowanej cząsteczki wody znajdującej się w centrum katalitycznym enzymu. Co ważne, postulowany mechanizm deaminacji (opis mechanizmu w podrozdziale 1.1.3.) [3] zakłada bezpośredni atak nukleofilowy tej cząsteczki na węgiel C4 pierścienia pirymidynowego. Bliskość silnie hydrofilowej grupy hydroksylowej i aktywowanej cząsteczki wody w miejscu aktywnym enzymu mogłaby skutkować powstaniem

wiązania wodorowego między nimi (Rysunek 4.5.2.D). W konsekwencji, atak nukleofilowy aktywowanej cząsteczki wody byłby niemożliwy i reakcja deaminacji nie mogłaby zajść. Na podstawie powyższych rozważań wysunięto wniosek, iż obserwowana niezdolność badanych wariantów hAID do deaminacji 5hmC wynika głównie z hydrofilowych właściwości grupy hydroksylowej, a nie – jak wcześniej postulowano – z zawady przestrzennej w centrum katalitycznym hAID wobec nieuwodnionej formy 5hmC.

Na podstawie przeprowadzonej analizy stwierdzono, że kieszenie katalityczne WT hAID i mutantu N51A są wystarczająco duże, aby pomieścić zarówno metylowaną jak i niemetylowaną cytydynę. Aby wzmocnić tą tezę dodatkowymi dowodami, zadokowano monobisfosforan cytydyny i monobisfosforan 5-metylocytydyny do centrów katalitycznych WT hAID i mutantu N51A, a następnie przeprowadzono 50 ns symulacji MD. Dokowanie substratów zostało przeprowadzone metodą dokowania ręcznego (ang. *manual docking*) we współpracy z dr. Karolem Kamelem z Zakładu Biologii Molekularnej i Systemowej ICHB PAN. Przy dokowaniu substratów uwzględniono postulowany mechanizm reakcji deaminacji cytydyny (opis mechanizmu w podrozdziale 1.1.3.). Ponadto podczas pozycjonowania cytydyny i 5-metylocytydyny uwzględniono następujące kryteria: (i) prawidłowe pozycjonowanie węgla C4 w stosunku do jonu cynku i aktywowanej cząsteczki wody (umożliwiającej zajście ataku hydrofilowego na ten atom); (ii) bliskość grupy karboksylowej reszty E58 do fragmentu aminowego pierścienia cytydyny lub 5-metylocytydyny; (iii) w przypadku WT hAID, bliskość łańcucha bocznego reszty N51 i grupy karbonylowej cytydyny lub 5-metylocytydyny.

Analiza trajektorii MD potwierdziła, że oba substraty mieszczą się w kieszeniach katalitycznych WT hAID i mutantu N51A. Oba substraty pozostały wewnątrz kieszeni katalitycznych w trakcie trwania symulacji, a w przypadku monobisfosforanu cytydyny można było zaobserwować pustą przestrzeń w okolicach węgla C5. Wynik ten, którego ilustracją są pliki video od 1 do 4 (patrz załącznik CD do niniejszej pracy), w sposób jednoznaczny potwierdził, że nie ma zawady przestrzennej w kieszeni katalitycznej WT hAID ani jej mutantu N51A wobec metylowanego substratu. Efekty steryczne nie były zatem czynnikiem warunkującym różnice w deaminacji C i 5mC przez te enzymy.



Rysunek 4.5.2. Proponowany mechanizm leżący u podstaw braku aktywności hAID wobec 5hmC.

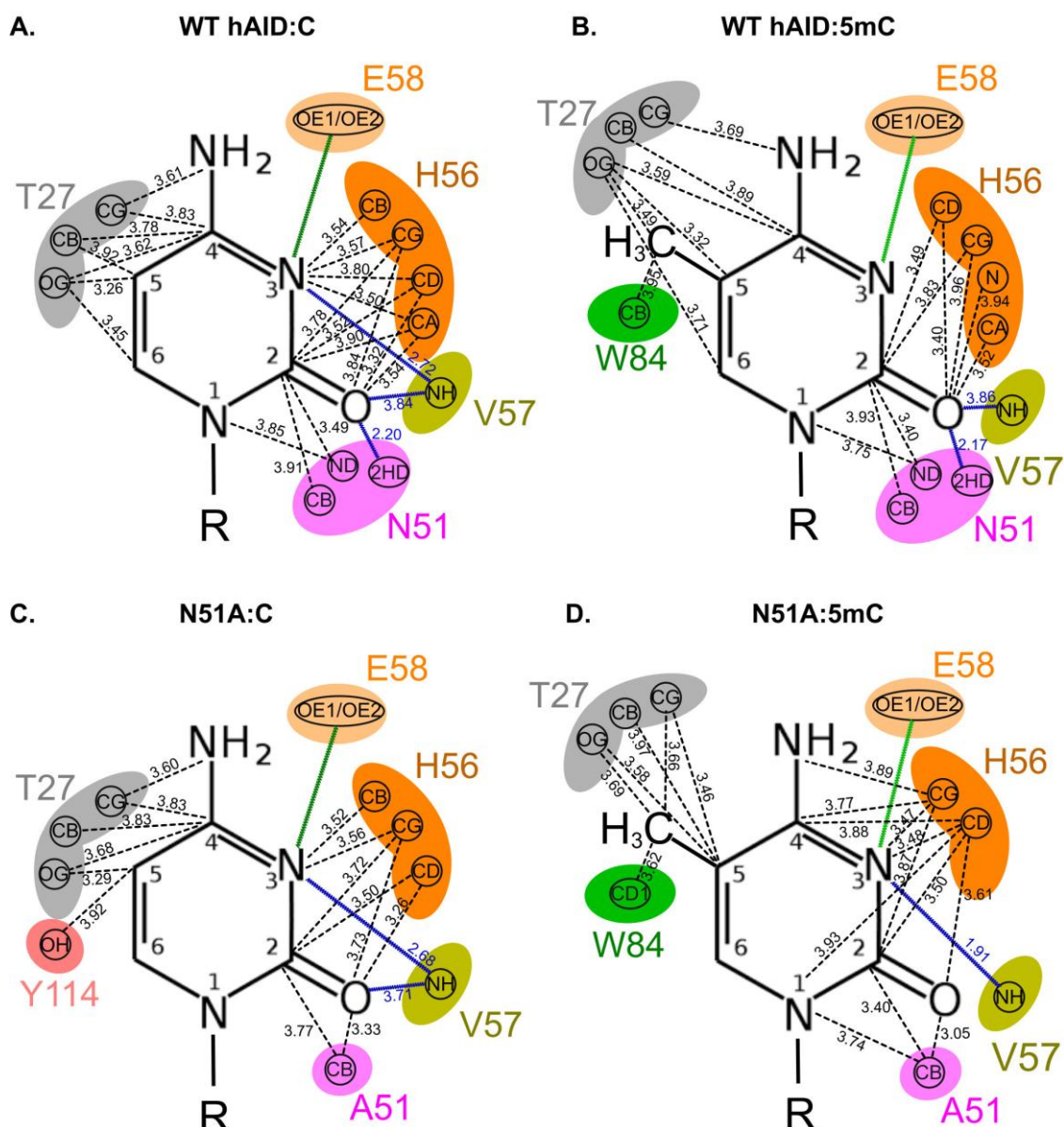
A) Kieszeń katalityczna hAID tworzona w znacznej części przez hydrofobowe reszty aminokwasowe (oznaczone kolorem pomarańczowym). **B)** Możliwa rotacja grupy hydroksylowej 5hmC. **C)** Postulowany mechanizm deaminacji niemodyfikowanego substratu. **D)** Proponowany mechanizm leżący u podstaw braku aktywności hAID wobec 5hmC. Niebieską elipsą zaznaczono wiązanie wodorowe mogące tworzyć się pomiędzy hydroksylem a grupą hydroksymetylową. Powstanie tego wiązania uniemożliwia atak nukleofilowy aktywowanej cząsteczki wody na węgiel C4. Nie dojdzie zatem do powstania stanu przejściowego (opis w tekście).

Poszukiwanie elementów strukturalnych warunkujących aktywność WT hAID i jej mutantu N51A wobec metylowanych i niemetylowanych substratów. Biorąc pod uwagę powyższe obserwacje, postawiono hipotezę, że zdolność WT hAID i mutantu N51A do deaminacji C i 5mC jest kształtowana przez zmiany w układzie oddziaływań między enzymem a substratem. Jeśli hipoteza ta byłaby prawdziwa, to mutacja N51A skutkowałaby utratą oddziaływań odpowiedzialnych za deaminację C ale zachowywałaby (lub tworzyłaby nowe) oddziaływania odpowiedzialne za deaminację 5mC.

Aby zidentyfikować czynniki warunkujące rozpoznawanie obu substratów wykonano dwie serie analiz. Pierwsza, skupiła się na określeniu sieci oddziaływań między pierścieniami pirymidynowymi C i 5mC a atomami tworzącymi kieszenie katalityczne WT hAID i mutantu N51A. Druga seria analiz koncentrowała się na identyfikacji oddziaływań pomiędzy WT hAID lub mutantem N51A a czteronukleotydowym ssDNA (AGCT lub AG5mCT).

Przetestowano cztery kompleksy typu enzym:substrat: WT hAID:C; WT hAID:5mC; N51A:C; N51A:5mC w trakcie symulacji MD modeli WT hAID i mutantu N51A z zadokowanym monobisfosforanem cytydyny lub monobisfosforanem 5-metylocytydyny (opis powyżej). Na tej podstawie zidentyfikowano i porównano wszystkie możliwe kontakty atomowe między pierścieniami C i 5mC, a atomami tworzącymi kieszenie katalityczne WT hAID i mutantu N51A. Jak przedstawiono na Rysunku 4.5.3. oraz w Załączniku 2 kompleks WT hAID:C charakteryzował się największą liczbą potencjalnych kontaktów – stwierdzono 24 możliwe kontakty (20 typu oddziaływań van der Waalsa (vdW) oraz 4 o charakterze wiązań wodorowych). Nieaktywny kompleks N51A:C charakteryzował się najmniejszą liczbą potencjalnych kontaktów – 17 (14 vdW i 3 wiązania wodorowe). Kompleksy WT hAID:5mC i N51A:5mC wykazywały podobną (odpowiednio 19 i 20), pośrednią liczbę potencjalnych kontaktów (16 vdW i 3 wiązania wodorowe dla kompleksu WT hAID:5mC oraz 18 vdW i 2 wiązania wodorowe dla kompleksu N51A:5mC). Obserwacje te sugerują, że gęstość kontaktów atomowych w obrębie tych dwóch kieszeni katalitycznych (o podobnych wymiarach) może być jednym z czynników decydujących o właściwym ułożeniu pierścienia C lub 5mC i ich stabilizacji w centrum katalitycznym. Wspólne dla wszystkich czterech testowanych kompleksów były oddziaływania pomiędzy: (i) pierścieniem cytozyny a resztą aminokwasową T27 (tworzącą „podstawę” kieszeni katalitycznej); (ii) pierścieniem cytozyny a resztą

aminokwasową H56 (koordynującą atom cynku w „sklepieniu” kieszeni katalitycznej); (ii) jedno lub dwa wiązania (o charakterze wiązań wodorowych) pomiędzy resztą aminokwasową V57 a regionem N3-C2 pierścienia cytozyny. Stwierdzono, że znaczącą różnicą pomiędzy testowanymi kompleksami jest brak, w wyniku mutacji N51A, możliwości tworzenia wiązań wodorowych pomiędzy resztą aminokwasową 51 a grupą karbonylową cytozyny, jak również brak trzech dodatkowych kontaktów typu vdW (z N1-C2 regionem pierścienia substratu). W kompleksach WT hAID:5mC i N51A:5mC obecność grupy metylowej skutkowała dodatkowymi kontaktami vdW pomiędzy grupą metylową, a resztami aminokwasowymi W84 i T27. Co ciekawe, analizy trajektorii MD wykazały, że reszta W84 tworzy, w testowanych kompleksach, niestabilne oddziaływanie warstwowe z resztą aminokwasową Y114. Co więcej dane literaturowe sugerują, że Y114 jest zaangażowana w rozpoznawanie czteronukleotydowego motywu, w obrębie którego niemodyfikowana cytydina ulega deaminacji [192]. Zidentyfikowane, niewielkie różnice pomiędzy kompleksami o skrajnie różnych właściwościach enzymatycznych wskazywały, że muszą istnieć jeszcze inne czynniki wpływające na rozpoznawanie i wiązanie C/5mC przez testowane warianty hAID.



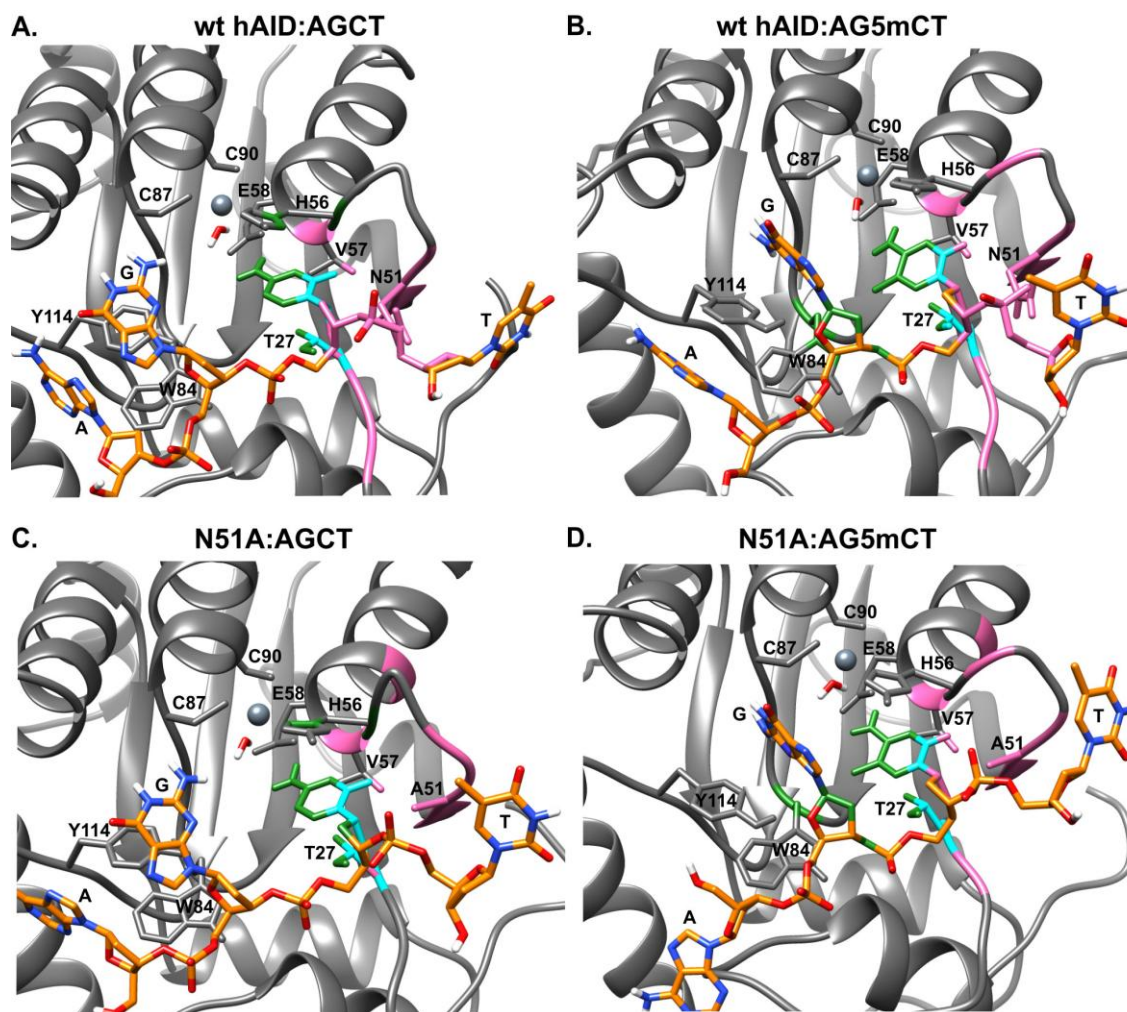
Rysunek 4.5.3. Kontakty atomowe stabilizujące pierścień C lub 5mC w obrębie czterech testowanych kompleksów enzym:substrat: (A) WT hAID:C; (B) WT hAID:5mC; (C) N51A:C; (D) N51A:C. Na rysunku różnymi kolorami zaznaczono poszczególne aminokwasy tworzące kieszeń katalityczną. W ich obrębie zaznaczono atomy zaangażowane w oddziaływanie z substratem. Czarne, przerywane linie oznaczają potencjalne kontakty vdW. Niebieskie linie – potencjalne wiązania wodorowe. Zieloną linią zaznaczono hipotetyczne wiązanie wodorowe, którego istnienie determinuje postulowany mechanizm deaminacji. Odległości pomiędzy poszczególnymi atomami kieszeni katalitycznej a atomami substratu podano w Å. Są to wartości średnie dla trajektorii MD.

W drugiej serii analiz, aby odtworzyć oddziaływania pomiędzy enzymem a ssDNA, zadokowano do kieszeni katalitycznych WT hAID i mutantu N51A czteronukleotydowe motywy AGCT i AG5mCT, po czym ponownie przeprowadzono 50 ns symulacji MD. Dokowanie czteronukleotydowych substratów zostało

przeprowadzone metodą dokowania ręcznego przez dr. Karola Kamela z Zakładu Biologii Molekularnej i Systemowej ICHB PAN. Przy dokowaniu substratów uwzględniono identyczne kryteria jak w przypadku dokowania monobisfosforanów. Analiza uzyskanych trajektorii potwierdziła, że reszta aminokwasowa N51 ma potencjał do tworzenia znacznie większej liczby oddziaływań niż alanina w tej pozycji (zarówno oddziaływań wewnątrzbiałkowych jak i oddziaływań z substratem) (Załącznik 3). Porównano potencjalne oddziaływania reszty N51 z substratami oraz A51 z substratami. Analiza ta wykazała, że zasięg potencjalnych oddziaływań N51 obejmował: region N1-C2 pierścienia cytozyny lub 5-metylocytozyny, wszystkie atomy deoksyrybozy cytydyny, grupę fosforanową oraz część deoksyrybozy tymidyny (T w obrębie motywu AGCT lub AG5mCT). Zasięg potencjalnych oddziaływań reszty A51 obejmował jedynie: region N1-C2 pierścienia cytozyny lub 5-metylocytozyny oraz C1'-O4' region deoksyrybozy cytydyny (Rysunek 4.5.4, Załącznik 3). Na tej podstawie stwierdzono, że najbardziej znaczącą różnicą pomiędzy aktywnym kompleksem WT hAID:AGCT a nieaktywnym N51A:AGCT jest wiązanie szkieletu cukrowo-fosforanowego ssDNA łączącego C lub 5mC oraz T w obrębie testowanych czternukleotydowych motywów. Co więcej resztę N51 różni od A51 możliwość tworzenia wiązań wodorowych z substratem (atomy HD21 i HD22 łańcucha bocznego N51 mogą tworzyć wiązania wodorowe z następującymi atomami substratów: O2, O4' oraz O3' cytydyny; O2P, O4' oraz O5' tymidyny; Załącznik 3)¹⁸. Przeprowadzone analizy oddziaływań sugerują, że reszta N51 jest kluczowa w wiązaniu szkieletu cukrowo-fosforanowego łączącego C i T w obrębie czternukleotydowego motywu AGCT. Co ważne, hAID nie deaminuje wolnej cytydyny. Fakt ten potwierdza obserwację, że wiązanie łańcucha cukrowo-fosforanowego ssDNA jest kluczowe dla aktywności enzymatycznej hAID.

Ponieważ mutant N51A wykazywał aktywność na metylovanym substracie (pomimo braku oddziaływań specyficznych dla reszty N51), w kolejnym etapie porównano możliwe kontakty atomowe grupy metylowej metylowanego substratu oraz węgla C5 niemetylowanej cytydyny. Jak pokazano na Rysunku 4.5.4 i w Załączniku 4 metylacja substratu skutkuje powstaniem dodatkowych kontaktów vdW pomiędzy grupą metylową a: resztą aminokwasową T27, resztą aminokwasową W84 oraz deoksyrybozą guaniny (G w obrębie motywu AG5mCT).

¹⁸ Oznaczenia poszczególnych atomów są zgodne z oznaczeniami atomów wygenerowanych modeli w programie Chimera 1.10.1, w którym wykonywane były analizy.



Rysunek 4.5.4. Porównanie zasięgu oddziaływań reszty N51 oraz A51, jak również metylu oraz węgla C5 w czterech kompleksach enzym:substrat: A) WT hAID:AGCT; B) WT hAID:AG5mCT; C) N51A:AGCT; D) N51A:AG5mCT. Różowym kolorem zaznaczono strefę oddziaływań reszty N51 lub A51 (w przypadku mutanta). Kolorem zielonym zaznaczono strefę oddziaływań grupy metylowej lub węgla C5 (w przypadku niemetylowanego substratu). Kolorem niebieskim zaznaczono obszar, gdzie obie powyższe strefy się nakładają. Reszta N51 oddziałuje z dużym fragmentem szkieletu cukrowo-fosforanowego ssDNA. Brak jest takich oddziaływań dla reszty A51. Widoczne jest również specyficzne tylko dla grupy metylowej oddziaływanie z resztą W84 oraz fragmentem szkieletu cukrowo-fosforanowego ssDNA łączącym G i C (w obrębie motywu AG5mCT). Oddziaływania warstwowe pomiędzy W84, Y114, a G są stabilne tylko w przypadku niemetylowanych substratów (opis w tekście).

Oddziaływanie grupy metylowej z jednej strony z białkiem, z drugiej, z deoksyrybozą guanidyny najprawdopodobniej było czynnikiem kompensującym brak oddziaływań tworzonych przez resztę N51 w kompleksie WT hAID:AGCT. Innymi słowy, zarówno grupa metylowa jak i reszta aminokwasowa N51 brały udział w wiązaniu szkieletu cukrowo-fosforanowego cząsteczki substratowej, w ten sposób zapewniając prawidłowe wiązanie substratu do centrum katalicznego hAID.

Analiza przeprowadzonych symulacji MD wykazała również, iż testowane kompleksy różniły się stabilnością oddziaływań warstwowych pomiędzy: W84, Y114 a nukleotydem G występującym w substracie. W przypadku kompleksów z niemetylowanym substratem: WT:AGCT i N51:AGCT, reszty W84, Y114 oraz G tworzyły trwałe oddziaływanie warstwowe (pliki video 5 i 7, załącznik CD do niniejszej pracy). W przypadku kompleksów z substratem metylovanym: WT:AG5mCT i N51:AG5mCT, obserwowano niestabilność tego układu, co w niektórych klatkach trajektorii skutkowało formowaniem się alternatywnych oddziaływań warstwowych pomiędzy G, resztą aminokwasową Y114 oraz A (pliki video 6 i 8, załącznik CD do niniejszej pracy). Na podstawie tej obserwacji można sądzić, iż obecność grupy metylowej wpływa na formowanie się oddziaływań warstwowych pomiędzy resztami aminokwasowymi W84 i Y114, a nukleotydami w pozycji -1 i -2 substratu. Należy jednak zauważyć, że zarówno aktywny kompleks WT:AGCT jak i nieaktywny N51A:AGCT charakteryzowały się trwałym układem oddziaływań warstwowych (W84 – Y114 – G), co sugeruje, że stabilność tego układu nie jest kluczowa dla aktywności enzymatycznej i prawdopodobnie pełni inną rolę. Wyraźny wpływ grupy metylowej na stabilność wymienionych wyżej oddziaływań warstwowych jest jednak kolejnym argumentem potwierdzającym oddziaływanie tej grupy ze szkieletem cukrowo-fosforanowym ssDNA.

5. DYSKUSJA

Otrzymywanie hAID w systemie prokariotycznym i jej wstępna charakterystyka.

Mimo że AID pełni bardzo istotne funkcje i cieszy się ogromnym zainteresowaniem badaczy w wiodących laboratoriach na świecie, enzym ten do niedawna nie był dostępny komercyjnie, a wiele pytań dotyczących jego aktywności enzymatycznej pozostaje bez jednoznacznej odpowiedzi. Podstawową przyczyną takiego stanu rzeczy jak i wielu kontrowersji, które narosły wokół aktywności AID, był brak wydajnego systemu do produkcji tego białka.

Aby scharakteryzować mechanizmy molekularne leżące u podstaw aktywności biologicznej hAID, w pierwszym etapie badań postawiono sobie za cel stworzenie wydajnego systemu umożliwiającego produkcję tego białka w formie aktywnej. Biorąc pod uwagę wcześniejsze obserwacje świadczące, iż modyfikacje potranslacyjne hAID *in vivo* nie wpływają znacząco na jej aktywność enzymatyczną, zdecydowano posłużyć się systemem bakteryjnym, ponieważ jest on najwydajniejszy i najtańszy. Opracowanie procedury produkcji hAID wiązało się z podjęciem wielu prób, w których testowano różne konstrukty i szczepy ekspresyjne a także warunki prowadzenia hodowli komórek bakteryjnych. Do otrzymania aktywnej formy białka niezbędna okazała się suplementacja pożywki LB (w której prowadzona była hodowla) chlorkiem cynku w stężeniu końcowym 0,06 mM. Skoordinowany w centrum katalitycznym hAID jon cynku (Zn^{2+}) jest niezbędny do zachowania aktywności enzymatycznej białka. Co ciekawe późniejsze dodanie chlorku cynku do preparatu uzyskanego z hodowli prowadzonej bez dodatku tej soli, nie prowadziło do uzyskania aktywnego białka. Na tej podstawie można przypuszczać, iż jon cynku jest bardzo silnie związany z hAID, oraz że jego obecność jest konieczna podczas biosyntezy białka. Nie można zatem wykluczyć możliwości, że Zn^{2+} wiąże się z hAID już podczas translacji i pozwala enzymowi przybrać aktywną konformację. Stąd w sytuacji braku Zn^{2+} powstaje nieaktywna forma białka. Innym ważnym czynnikiem wpływającym na wydajność produkcji aktywnego białka była temperatura. Od momentu indukcji ekspresji heterologicznego genu hodowlę prowadzono w temperaturze 18°C. Wydaje się, że obniżenie temperatury ogranicza powstawanie ciałek inkluzyjnych co jest szczególnie istotne w przypadku białek szkodliwych dla komórki (potencjał mutageny hAID powoduje, że należy ona do takich białek). Zastosowana metoda oczyszczania

pozwoiliła na uzyskanie preparatu o możliwie dużej homogenności przy jednoczesnym zachowaniu właściwości enzymatycznych produkowanego białka.

System bakteryjny był już wcześniej stosowany do otrzymywania AID [1, 193, 194]. Na podstawie danych literaturowych można jednak sądzić, że uzyskane w ten sposób preparaty były zanieczyszczone deaminazą bakteryjną [3]. Stąd w szeregu publikacjach pojawiły się sprzeczne informacje dotyczące właściwości AID, m.in. jej zdolności do deaminacji wolnego nukleozydu [1] czy ulegania inhibicji w obecności THU [195]. W celu wykluczenia możliwości, iż obserwowana aktywność wynika z obecności enzymu bakteryjnego przeprowadzono szereg analiz. Po pierwsze, potwierdzono metodami MALDI TOF i *western blot*, że jednym z podstawowych składników uzyskanego preparatu jest białko hAID. Po drugie wykazano, że aktywność otrzymanego białka jest inhibowana przez 1,10-fenantrolinę co dowiodło, że jest ono metaloproteiną. Po trzecie wykazano, że THU nie jest inhibitorem otrzymanego białka (bakteryjne deaminazy wolnej cytydyny są inhibowane przez THU). Po czwarte, przeprowadzono szereg testów kontrolnych na wszystkich etapach procedury badania aktywności hAID. Stwierdzono na ich podstawie, że obserwowana aktywność jest specyficzna jedynie wobec cząsteczek ssDNA zawierających C (nie stwierdzono aktywności wobec cząsteczek ssDNA niezawierających C). Wykluczono również potencjalne zanieczyszczenie preparatu nukleazami. Po piąte, uzyskano trzy mutanty hAID wykazujące różną aktywność wobec C, przy czym stwierdzona aktywność była zgodna z danymi literaturowymi. Podsumowując, przeprowadzone testy dowiodły, że aktywność deaminazowa otrzymanego preparatu białkowego jest aktywnością hAID, i że charakterystyka biochemiczna tej aktywności jest spójna z najnowszymi danymi literaturowymi.

W wyniku wyżej opisanych doświadczeń udało się uzyskać preparat hAID nie w pełni homogeny ale o zachowanej aktywności enzymatycznej. Opracowana metoda otrzymywania hAID i uzyskany w wyniku jej zastosowania enzym stały się przedmiotem zgłoszenia patentowego w Urzędzie Patentowym RP (o numerze: P402440, 2013), zgłoszenia międzynarodowego (o numerze: PCT/PL2014/000003, 2014), zgłoszenia europejskiego (o numerze: EP-14703938.2, 2014) oraz zgłoszenia w USA (o numerze: US 14/384,665, 2014). Co więcej, została również podpisana umowa (ang. *material transfer agreement*, MTA) między IChB PAN a laboratorium profesora Fredericka Alta z Harvard Medical School (jeden ze światowych liderów badań AID), na mocy której zarówno uzyskany konstrukt ekspresyjny, jak i opisana

powyżej metoda produkcji hAID zostały udostępnione temu laboratorium. Świadczy to o tym, jak dużym wyzwaniem było uzyskanie preparatu zawierającego aktywną hAID i jak dużym zainteresowaniem cieszy się ten preparat. W dniu 15 października 2014 roku została podpisana umowa licencyjna pomiędzy Instytutem Chemii Bioorganicznej PAN a firmą Genesis sp. z o.o. dotycząca komercjalizacji preparatu wykazującego aktywność hAID. W rezultacie, preparat zawierający aktywne białko hAID znalazł się w sprzedaży i aktualnie jest jedynym na świecie komercyjnym źródłem tego białka. Dnia 08 stycznia 2015 roku wspólne przedsięwzięcie firmy Genesis i IChB PAN jako jedno z trzech w Wielkopolsce zostało nominowane do prestiżowej Nagrody Marszałka Województwa Wielkopolskiego „i-Wielkopolska – Innowacyjni dla Wielkopolski” w kategorii „Innowacyjna Inwencja”. Uzyskany preparat wykorzystano w dalszych badaniach prowadzonych w ramach niniejszej pracy doktorskiej.

Aktywność hAID wobec modyfikowanych i niemodyfikowanych substratów.

Procesy biochemiczne leżące u podstawy aktywnej demetylacji genomu kręgowców nie są znane. Proponowane są trzy możliwe mechanizmy. Każdy z nich to wieloetapowy szlak, który polega na modyfikacji 5mC do innych związków oraz angażuje mechanizmy naprawy DNA [109, 127, 128]. Deaminazy z rodziny AID/APOBEC są ważnymi elementami dwóch z postulowanych szlaków. W pierwszym deaminazy AID/APOBEC miałyby bezpośrednio deaminować 5mC do T, co generowałoby powstanie niesparowania T-G, rozpoznawanego następnie przez system naprawy BER. W drugim, 5mC podlegałyby modyfikacji do 5hmC przy udziale białka TET1¹⁹. Udział deaminaz AID/APOBEC polegałby wówczas na deaminacji 5hmC do 5hmU. Produkt deaminacji również w tym przypadku byłby rozpoznawany przez maszynę BER [108, 110]. Potencjalny udział AID w tych procesach jest sugerowany od 2004 roku, kiedy to po raz pierwszy pokazano zdolność AID do katalizowania reakcji deaminacji 5mC do T, jak również stwierdzono ekspresję genu AID poza układem immunologicznym [55]. Od tamtej pory pojawiło się szereg doniesień potwierdzających udział AID w procesach demetylacji, jak i wiele prac poddających go w wątpliwość. Najwięcej zastrzeżeń budzi obniżone powinowactwo centrum aktywnego AID wobec 5mC w porównaniu z niemodyfikowanym substratem. W niektórych badaniach biochemicznych AID, wykazano że białko to posiada nawet 16-krotnie niższą

¹⁹ TET1 katalizuje reakcję oksydacji 5mC do 5hmC.

aktywność *in vitro* względem 5mC niż względem niemodyfikowanej cytydyny [123, 137]. W przypadku 5hmC istnieją dane *in vivo* przemawiające za udziałem AID w deaminacji tego związku [122], ale testy *in vitro* nie potwierdzają takiej możliwości [123, 124]. Część badaczy postuluje, że AID mniej wydajnie deaminuje 5mC i nie deaminuje 5hmC ze względu na zawadę przestrzenną wobec podstawnika przy węglu 5 występującą w centrum aktywnym tego enzymu [124, 138]. W rezultacie uwarunkowania steryczne postrzegane są jako znacząca bariera dla udziału deaminaz AID/APOBEC w procesie aktywnej demetylacji.

W obliczu powyższych wątpliwości, w badaniach podjętych w ramach niniejszej rozprawy doktorskiej postawiono szczegółowo scharakteryzować uwarunkowania strukturalne aktywności hAID wobec C, 5mC oraz 5hmC. W pierwszym etapie badań przetestowano aktywność deaminazową WT hAID jak i trzech jej mutantów wobec 80-nukleotydowych cząsteczek ssDNA zawierających tylko jedną resztę C, 5mC lub 5hmC ulokowaną w obrębie motywu AGCT, AG5mCT lub AG5hmCT. Spośród wielu scharakteryzowanych mutantów AID (Tabela 1.4.1) wybrano trzy: R50A, N51A, Δ 189. W przypadku R50A i N51A pojedyncza substytucja aminokwasowa dotyczyła reszt w pobliżu centrum katalicznego hAID. Mutant Δ 189 był wariantem białka skróconym o dziewięć reszt aminokwasowych z końca C. Wybrane mutanty, zgodnie z danymi literaturowymi, wykazywały różną aktywność *in vitro* wobec C (resztkową – mutant R50A, brak aktywności – mutant N51A, wyższą niż WT hAID – mutant Δ 189). Co ciekawe, zgodnie z danymi literaturowymi aktywność *in vitro* tych mutantów nie korelowała z wydajnością co najmniej jednego z procesów inicjowanych przez AID (SHM lub CSR) w warunkach *in vivo* (Tabela 4.3.1). Ponadto, aktywność deaminazowa powyższych mutantów hAID na modyfikowanych substratach nie była wcześniej badana. W przeprowadzonych w ramach niniejszej pracy testach *in vitro*, żaden z badanych wariantów hAID nie wykazywał aktywności wobec 5hmC. Wszystkie testowane warianty hAID wykazywały natomiast zróżnicowaną aktywność wobec C i 5mC. Co ciekawe aktywności testowanych mutantów wobec C i 5mC nie korelowały ze sobą (każda z mutacji różnie wpływała na zdolność hAID do deaminacji C i 5mC). Oznacza to, że poprzez mutagenezę można, w sposób niezależny, modyfikować specyficzność hAID wobec metylowanego i niemetylowanego substratu. Doskonałym przykładem ilustrującym przedstawione powyżej obserwacje jest odmienna aktywność deaminazowa stwierdzona dla dwóch wariantów hAID: enzymu typu dzikiego oraz

mutanta N51A. Pierwszy jest aktywny wobec C i 5mC, drugi jest nieaktywny wobec C i aktywny wobec 5mC.

W kolejnym etapie przeprowadzonych badań potwierdzono całkowity brak aktywności mutantu N51A wobec niemetylowanych C występujących w ssDNA. Użyte w doświadczeniach cząsteczki ssDNA zawierały liczne C w obrębie wszystkich możliwych motywów sekwencyjnych typu NNCN (gdzie N oznacza dowolny nukleotyd). Mutant N51A nie wykazywał aktywności ani wobec preferowanych przez WT hAID motywów sekwencyjnych (WRCY), ani żadnych innych. W przeciwieństwie do mutantu, WT hAID wykazywała aktywność wobec C znajdującej się w obrębie różnych motywów sekwencyjnych. Dla WT hAID obserwowano zarówno cząsteczki ssDNA, w których wszystkie lub większość C było zdeaminowanych, jak i takie, w których żadna C nie była zdeaminowana. Obserwacja ta potwierdza hipotezę postulowaną przez grupę profesora Goodmana, że aktywność AID ma charakter procesywny.

Mutant N51A mysiej AID był wcześniej badany przez grupę profesora Honjo [35]. Badania tej grupy pokazały, że mutant ten nie deaminuje C w warunkach *in vitro*. Co zaskakujące, komórki posiadające tą mutację w genie AID wykazywały 50% wydajność procesu CSR w porównaniu z WT AID. Na podstawie obserwacji, że proces CSR zachodzi w komórkach z mutacją N51A badacze wysunęli wniosek, iż aktywność deaminazowa AID nie reprezentuje fizjologicznej funkcji tego białka. W badaniach tych nie sprawdzono jednak możliwości, iż mutant N51A rozpoznaje modyfikowaną cytydynę. Wyniki zawarte w niniejszej rozprawie doktorskiej oraz wyniki grupy profesora Honjo łącznie sugerują, iż mutant N51A może deaminować 5mC w *locus Ig* w trakcie procesu CSR. Konsekwentnie, deaminacja 5mC może potencjalnie zachodzić również w innych procesach biologicznych takich jak różnicowanie przeciwciał, a nie jak do tej pory sądzono, jedynie w procesach związanych z aktywną demetylacją DNA.

AID działa na ssDNA zatem jej zdolność do deaminacji w warunkach *in vivo* jest uzależniona od aktywności transkrypcyjnej rozpoznawanych regionów genomu. Metylacja DNA jest natomiast najczęściej kojarzona z represją transkrypcji. Ta sprzeczność może być uznana za argument przemawiający przeciwko udziałowi AID w deaminacji 5mC *in vivo*. Najnowsze dane literaturowe pokazują jednak, iż metylacja DNA może występować również wewnątrz aktywnych transkrypcyjnie genów. Znaczenie i funkcja takiej wewnątrzgenowej metylacji DNA aktywnych transkrypcyjnie

genów pozostaje nieznana [140]. Jej istnienie wspiera jednak hipotezę, iż AID może deaminować 5mC *in vivo*.

Opisane powyżej wnioski, dotyczące właściwości enzymatycznych mutantu N51A, znajdują poparcie w doniesieniach literaturowych dotyczących właściwości i struktury białek spokrewnionych z AID. Po pierwsze, reszta asparaginowa (wymieniona w N51A na alaninę) jest wysoce konserwatywna w obrębie ortologów AID [35]. Po drugie, substytucja alaniną asparaginy N244 w białku APOBEC3G (reszta N244 w APOBEC3G jest odpowiednikiem reszty N51 w AID) znosi aktywność deaminazową tego białka wobec C. Aktywność mutantu N244A białka APOBEC3G wobec 5mC niestety nie była badana. Po trzecie, asparagina N51 w ortologach AID może być zaangażowana w wiązanie substratu. Posługując się metodami krystalograficznymi określono, że odpowiadająca reszcie N51 reszta aminokwasowa w deaminazie adenozyliny (działającej na tRNA) oddziałuje z substratem poprzez wiązania wodorowe. Na tej podstawie zaproponowano oddziaływanie reszty N51 z C również w przypadku AID [6, 30, 31].

W kolejnym etapie przeprowadzonych badań przeanalizowano aktywność testowanych wariantów hAID wobec 5mC w kontekście genomowym, wobec długich odcinków ssDNA zawierających liczne 5mC. Uzyskane wyniki pokazały, że zarówno WT hAID jak i jej mutanty: R50A, N51A, Δ 189 w warunkach *in vitro* drastycznie zmniejszają globalny poziom 5mC (redukując go o 70 do 90%). Na tej podstawie wysunięto wniosek, że hAID jest zdolna do obniżania globalnego poziomu metylacji gDNA *in vitro* bez udziału żadnych kofaktorów, oraz że hAID może deaminować 5mC w długich odcinkach ssDNA zawierających liczne 5mC. W warunkach *in vitro* wpływ AID na globalny poziom metylacji nigdy wcześniej nie był badany.

Pokazana w powyższych badaniach zdolność mutantu N51A do rozróżniania pomiędzy metylovaną i niemetylovaną cytydyną w długich odcinkach ssDNA może znaleźć praktyczne implikacje np. jako alternatywa dla reakcji konwersji wodorosiarczynem sodu w badaniach profilu metylacji DNA. W analizach metylacji DNA na drodze sekwencjonowania DNA traktowanego wodorosiarczynem sodu (w tzw. sekwencjonowaniu wodorosiarczynowym), związek ten służy do konwersji niemodyfikowanej cytydyny do tymidyny (metylowana cytydyna pozostaje niezmienną). Biorąc pod uwagę: (i) selektywność działania mutantu N51A wobec 5mC; (ii) potencjalnie procesywny charakter tej aktywności; (iii) stwierdzoną, na podstawie wyników niniejszych badań, zdolność tego mutantu do aż 80% redukcji

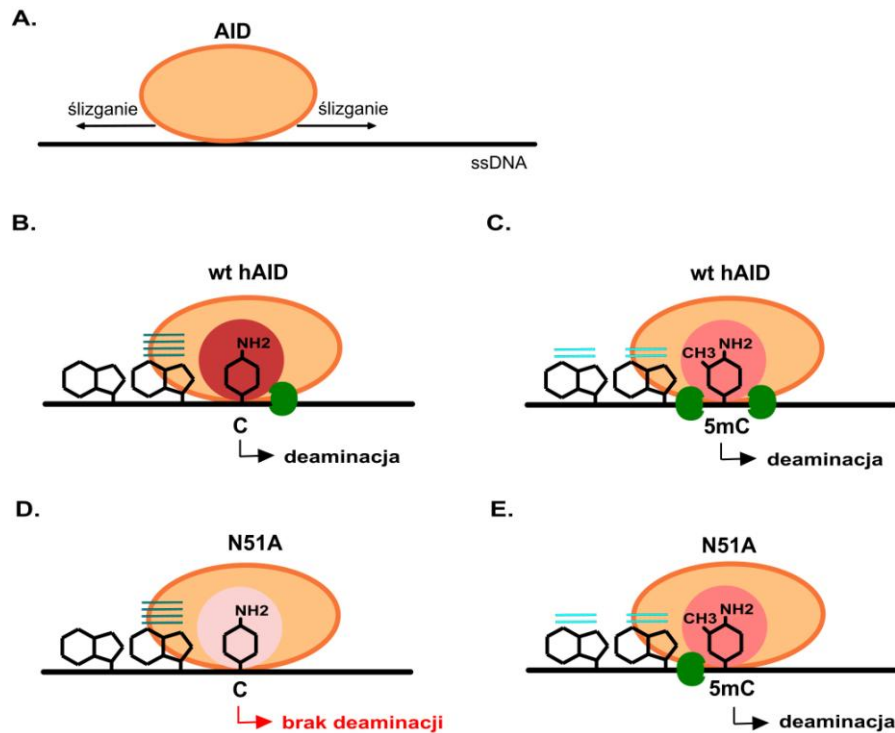
poziomu globalnej metylacji gDNA *in vitro*; wydaje się możliwe jego praktyczne zastosowanie w konwersji odwrotnej niż konwersja wodorosiarczynowa (a więc metylowanej cytydyny do tymidyny). Wyniki przeprowadzonych badań mają zatem również aspekt praktyczny.

Obserwowany brak korelacji między aktywnością testowanych mutantów hAID wobec C i 5mC wskazuje, iż muszą istnieć znaczące różnice w molekularnych mechanizmach rozpoznawania metylowanych i niemetylowanych substratów przez hAID. Dane literaturowe sugerują istnienie zawady przestrzennej w kieszeni katalitycznej AID wobec podstawnika przy węglu 5 cytydyny mającej ulec deaminacji [123, 124]. Na podstawie wyników modelowania homologicznego i dynamiki molekularnej wykazano w niniejszej pracy, iż zarówno w kieszeni katalitycznej WT hAID jak i jej mutantu N51A jest wystarczająco dużo miejsca na związanie C lub 5mC. Nie stwierdzono zatem występowania zawady przestrzennej w centrum katalitycznym wobec tych substratów. W dalszych analizach *in silico* kompleksów hAID:substrat zidentyfikowano trzy elementy, które determinują czy C lub 5mC może ulec deaminacji. Pierwszym elementem jest sieć oddziaływań pomiędzy kieszenią katalityczną a pierścieniem pirymidynowym C lub 5mC. Oddziaływania te stabilizują substrat w pozycji katalitycznej ale nie są wystarczające do tego aby zaszła reakcja deaminacji (co tłumaczy dlaczego hAID nie deaminuje wolnych nukleozydów). Drugim, zidentyfikowanym elementem jest sieć oddziaływań pomiędzy hAID a szkieletem cukrowo-fosforanowym ssDNA. W przypadku substratów zawierających C, sieć tą budują oddziaływania pomiędzy resztą N51 a szkieletem cukrowo-fosforanowym łączącym C i T w obrębie motywu AGCT. Oddziaływania te są niezbędne aby zaszła deaminacja C. Brak kontaktów atomowych tworzonych przez N51 skutkowało brakiem aktywności enzymatycznej mutantu N51A wobec niemetylowanego substratu. W przypadku substratów zawierających 5mC, grupa metylowa pełni rolę łącznika oddziałując, z jednej strony z hAID (T27, W84), a z drugiej strony ze szkieletem cukrowo-fosforanowym ssDNA (reszta cukrowa G w obrębie motywu AGCT). Oddziaływania tworzone przez grupę metylową są wystarczające dla prawidłowego wiązania metylowanego substratu, nawet przy braku oddziaływań tworzonych przez resztę N51. Tłumaczy to wyniki eksperymentalne zawarte w niniejszej pracy, w których to metylacja substratu kompensowała brak oddziaływań tworzonych przez resztę N51 i przywracała aktywność mutantu N51A.

Trzecim zidentyfikowanym elementem są oddziaływania warstwowe pomiędzy: W84, Y114 a G poprzedzający C w obrębie motywu AGCT. Zgodnie z najnowszymi danymi literaturowymi oddziaływania warstwowe tworzone przez Y114 uczestniczą w rozpoznawaniu kontekstu sekwencyjnego cytydyny ulegającej deaminacji [192]. Analizy przeprowadzone w niniejszej pracy pokazały, że obecność grupy metylowej destabilizuje ten układ oddziaływań warstwowych. W rezultacie w kompleksach hAID z metylowanym substratem układ ten nie występuje lub jest zastępowany przez alterantywne oddziaływania warstwowe pomiędzy A, Y114 i G (A i G w obrębie motywu AGCT). Obserwacja ta sugeruje, że preferowany przez hAID motyw sekwencyjny, w obrębie którego cytydyna ulega deaminacji może być różny dla metylowanych i niemetylowanych substratów. Aby to potwierdzić konieczne jest przeprowadzenie dodatkowych badań np. aktywność WT hAID i jej mutantów na metylowanych, różnych czteronukleotydowych motywach.

Podsumowując, każdy z trzech wyżej zidentyfikowanych elementów (składających się na sieć oddziaływań hAID:substrat) jest różny w przypadku motywu AGCT i AG5mCT. Wskazuje to, że oba te substraty są odmiennie rozpoznawane przez hAID.

Na podstawie danych literaturowych można sądzić, iż hAID działa procesywnie. Zgodnie z mechanizmem zaproponowanym przez Chelico i współautorów [186], hAID wiąże się do ssDNA i ślizga wzdłuż cząsteczki w obu kierunkach. Jedynie okazjonalnie enzym zatrzymuje się by zdeaminować C. Zaproponowany mechanizm nie tłumaczy jednak jakie czynniki muszą być spełnione aby deaminacja zaszła. Wyniki badań przeprowadzonych w ramach niniejszej pracy są zgodne z hipotezą Chelico i współautorów, a dodatkowo sugerują, iż wszystkie trzy wyżej wspomniane elementy decydują o skłonności enzymu do zatrzymania się i deaminacji C lub 5mC. Na podstawie wyników niniejszych badań zaproponowano następujący scenariusz zależnej od hAID deaminacji (Rysunek 5.1.1). Dochodzi do niej gdy hAID zatrzyma się (zakotwiczy) na C lub 5mC w trakcie ślizgania się wzdłuż ssDNA. Głównym elementem odpowiedzialnym za kotwiczenie się hAID do substratu jest sieć oddziaływań pomiędzy białkiem a szkieletem cukrowo-fosforanowym ssDNA.



Rysunek 5.1.1. Mechanizm zależnej od hAID deaminacji C lub 5mC. A) hAID wiąże się do ssDNA i ślizga wzdłuż cząsteczki w obu kierunkach. B-E) Enzym zatrzymuje się okazjonalnie i deaminuje C lub 5mC. Trzy elementy determinują skłonność enzymu do deaminacji: (1) sieć oddziaływań pomiędzy kieszenią katalityczną enzymu a pierścieniem pirymidynowym (różowe koło, odcień koloru odpowiada sile oddziaływań – im ciemniejszy kolor tym silniejsze oddziaływania); (2) sieć oddziaływań pomiędzy enzymem a szkieletem cukrowo-fosforanowym ssDNA (zielone koło); (3) oddziaływania warstwowe pomiędzy resztami aromatycznymi w białku a zasadą azotową (lub zasadami azotowymi) sasiadującą z C lub 5mC (niebieskie linie, odcień koloru i liczba linii odpowiadają stabilności oddziaływań – im ciemniejszy kolor i więcej linii tym stabilniejsze oddziaływania). **B)** Zdolność WT hAID do kotwiczenia się i deaminacji C jest warunkowana przez: (1) stabilizację pierścienia pirymidynowego w centrum katalitycznym; (2) oddziaływania reszty N51 ze szkieletem cukrowo-fosforanowym ssDNA; (3) stabilny układ oddziaływań warstwowych (pomiędzy W84, Y114 i G). **C)** Zdolność WT hAID do kotwiczenia się i deaminacji 5mC jest warunkowana przez: (1) wystarczającą stabilizację pierścienia pirymidynowego w centrum katalitycznym; (2) podwójne kotwiczenie szkieletu cukrowo-fosforanowego ssDNA (poprzez oddziaływania specyficzne dla reszty N51 i grupy metylowej); (3) niestabilny układ oddziaływań warstwowych (pomiędzy W84, Y114 i G) lub alternatywny układ oddziaływań warstwowych (pomiędzy A, Y114 i G). **D)** Na brak zdolności mutantu N51A do deaminacji C składają się: (1) niewystarczająca stabilizacja pierścienia pirymidynowego w centrum katalitycznym; (2) brak kotwiczenia szkieletu cukrowo-fosforanowego ssDNA. W tym przypadku jedynie układ oddziaływań warstwowych pomiędzy W84, Y114 i G jest stabilny. **E)** Mutant N51A ma zdolność deaminacji 5mC dzięki: (1) wystarczającej stabilizacji pierścienia pirymidynowego w centrum katalitycznym; (2) kotwiczeniu szkieletu cukrowo-fosforanowego ssDNA (poprzez oddziaływania specyficzne dla grupy metylowej); (3) niestabilnemu układowi oddziaływań warstwowych (pomiędzy W84, Y114 i G) lub alternatywnemu układowi oddziaływań warstwowych (pomiędzy A, Y114 i G).

Brak tych oddziaływań w kompleksie N51A:AGCT skutkuje brakiem aktywności enzymu wobec C. Dodatkowym elementem wzmacniającym efekt kotwiczenia są oddziaływania warstwowe pomiędzy W84, Y114 a nukleotydem poprzedzającym C. Oddziaływania warstwowe pomiędzy aminokwasami aromatycznymi, a pierścieniami nukleotydów dodatkowo wpływają na specyficzność działania hAID w obrębie motywów sekwencyjnych. Trzecim elementem jest stabilizacja pierścienia pirymidynowego C lub 5mC w kieszeni katalitycznej enzymu. W kompleksach WT hAID:AGCT, WT hAID:AG5mCT, N51A:AG5mCT oddziaływania kieszeni katalitycznej z pierścieniem pirymidynowym są wystarczające dla zapewnienia pozycji katalitycznej. W przypadku kompleksu N51A:AGCT sieć tych oddziaływań jest uboższa co prawdopodobnie przyczynia się do braku aktywności enzymatycznej.

Na podstawie badań przeprowadzonych w niniejszej pracy można sądzić, iż metylowany substrat jest podwójnie kotwiczony w centrum katalitycznym WT hAID (szkielet cukrowo-fosforanowy ssDNA jest wiązany zarówno od strony grupy metylowej jak i od strony reszty N51). Niemetylowany substrat jest natomiast wiązany tylko jednostronnie (przez resztę N51). Podwójne kotwiczenie ssDNA przez enzym, hipotetycznie może skutkować: (i) wzmocnieniem efektu zatrzymywania się (kotwiczenia) enzymu podczas jego ślizgania się wzdłuż ssDNA, a w rezultacie wzrostem efektywności deaminacji; (ii) spadkiem szybkości oddysocjowywania produktu reakcji, a zatem spadkiem szybkości katalizy. Druga z możliwości wydaje się dominować biorąc pod uwagę dane literaturowe wskazujące na 3-, 10- lub 16-krotny spadek aktywności WT AID wobec 5mC w porównaniu z jej aktywnością wobec C w warunkach *in vitro* [32, 55, 123]. Podwójne kotwiczenie metylowanego substratu może być zatem przyczyną obserwowanego spadku efektywności działania AID wobec 5mC w porównaniu z C. W niniejsze pracy nie porównano bezpośrednio aktywności WT hAID wobec C i 5mC. Nie było to możliwe ze względu na to, że metody detekcji produktów deaminacji tych dwóch substratów były inne. Opisywany w literaturze spadek aktywności AID na 5mC, część badaczy uznaje za argument przemawiający przeciwko jej zaangażowaniu w procesy demetylacji. Jednakże, taki spadek aktywności ostatecznie nie wyklucza możliwości deaminowania przez AID 5mC *in vivo*. Trzeba bowiem pamiętać, iż AID deaminuje preferowany motyw WRCY (ang. *hot spot*) od 4 do 8 razy bardziej efektywnie niż motywy niepreferowane (ang. *cold spot*) przy czym deaminacja niepreferowanych motywów jest również wykrywana w warunkach *in vivo* w trakcie procesu SHM [187]. W tym kontekście różna wydajność deaminacji

metylowanych i niemetylowanych substratów przestaje być czynnikiem, który wyklucza możliwość zależnej od AID deaminacji 5mC *in vivo* w trakcie różnych (może jeszcze niepoznanych) procesów biologicznych.

W przeprowadzonych badaniach żaden z testowanych wariantów hAID nie wykazywał wykrywalnej aktywności deaminazowej wobec 5hmC. Ten wynik jest zgodny z rezultatami badań *in vitro* innych grup badawczych [123, 124]. Zaproponowane w niniejszej pracy wyjaśnienie braku aktywności hAID wobec 5hmC jest jednak zasadniczo różne od postulowanego wcześniej. Zakłada ono bowiem, że głównym czynnikiem uniemożliwiającym deaminację 5hmC jest silnie hydrofilowy charakter grupy hydroksylowej. Powoduje on: (i) hydratację tej grupy w roztworach wodnych, przez co pierścień 5hmC staje się zbyt duży aby wejść do kieszeni katalitycznej hAID (dodatkową przeszkodą jest fakt, iż ta kieszeń jest tworzona w dużej części przez aminokwasy o charakterze hydrofobowym); (ii) zaburzenie mechanizmu deaminacji w przypadku gdyby dehydratowana 5hmC znalazłaby się w kieszeni katalitycznej (grupa hydroksylowa ma potencjał do stworzenia wiązania wodorowego z aktywowaną cząsteczką wody znajdującą się w centrum aktywnym hAID).

Wyniki zaprezentowane w niniejszej pracy skłaniają do ponownego rozważenia dotychczasowych poglądów, w myśl których 5mC jest postrzegana jako gorszy substrat dla hAID. Opisane rezultaty pozwalają sądzić, że hAID wykazuje różną specyficzność wobec metylowanych i niemetylowanych substratów. Taka podwójna specyficzność wobec metylowanego i niemetylowanego DNA została w ostatnim czasie pokazana dla niektórych czynników transkrypcyjnych [196] i być może, w niedalekiej przyszłości, okaże się zjawiskiem dość powszechnym w przypadku białek wiążących DNA. Ponadto, uzyskane wyniki już dziś uzasadniają twierdzenie, że udział AID w procesie aktywnej demetylacji, jak i innych procesach biologicznych związanych z deaminacją 5mC jest możliwy ze względu na właściwości enzymatyczne tego białka. Jednocześnie wyniki te wspierają hipotezę, iż udział AID w procesach demetylacji DNA na drodze deaminacji 5hmC jest mało prawdopodobny.

6. WNIOSKI

Badania przeprowadzone w ramach niniejszej pracy prowadzą do następujących wniosków:

1. Uzyskanie aktywnego preparatu hAID jest możliwe w systemie prokariotycznym pod warunkiem prowadzenia hodowli w pożywce suplementowanej chlorkiem cynku oraz w temperaturze 18°C (od momentu indukcji ekspresji heterologicznego genu).
2. Aktywność enzymatyczna otrzymanego białka jest hamowana przez 1,10-fenantrolinę, jest specyficzna jedynie wobec cząsteczek ssDNA zawierających C lub 5mC, oraz nie jest hamowana przez THU.
3. hAID i jej mutanty nie wykazują aktywności wobec 5hmC *in vitro*.
4. hAID i jej mutanty wykazują *in vitro* różną aktywność wobec C i 5mC, przy czym aktywności wobec tych dwóch substratów nie korelują ze sobą, co świadczy o istnieniu różnych uwarunkowań strukturalnych deaminacji metylowanych i niemetylowanych substratów.
5. WT hAID działa w sposób procesywny.
6. Zarówno WT hAID jak i jej mutanty mają zdolność do obniżenia globalnego poziomu 5mC w obrębie gDNA oraz deaminacji 5mC w długich odcinkach ssDNA zawierających liczne 5mC.
7. Mutant N51A, nieaktywny wobec C, wykazuje aktywność wobec 5mC. Mutant ten może znaleźć praktyczne zastosowanie np. w badaniach metylacji DNA.
8. W centrum katalitycznym hAID nie występuje zawada przestrzenna wobec 5mC.
9. Brak aktywności hAID wobec 5hmC może być tłumaczony właściwościami chemicznymi grupy hydroksylowej i wpływem tych właściwości na przebieg reakcji deaminacji.
10. Metylowane i niemetylowane substraty są różnie wiązane przez kieszeń katalityczną hAID.
11. Metylacja substratu kompensuje brak oddziaływań specyficznych dla reszty N51 dzięki czemu mutant N51A wykazuje aktywność wobec 5mC przy braku aktywności wobec C.
12. Grupa metylowa wpływa na układ oddziaływań warstwowych tworzonych przez resztę Y114. Według danych literaturowych układ ten jest odpowiedzialny za rozpoznawanie kontekstu sekwencyjnego C ulegającej deaminacji. Sugeruje to, że hAID może wykazywać różną specyficzność wobec metylowanych

i niemetylowanych motywów sekwencyjnych. Zagadnienie to powinno zostać szczegółowo zbadane w przyszłości.

13. Podawana w literaturze niższa aktywność hAID wobec 5mC w porównaniu z C może być tłumaczona podwójnym kotwiczeniem modyfikowanego substratu w kieszeni katalitycznej tego enzymu.
14. Udział hAID w procesach aktywnej demetylacji genomu, jak i innych procesach związanych z deaminacją 5mC, jest możliwy ze względu na właściwości enzymatyczne tego białka.
15. Udział hAID w procesach aktywnej demetylacji genomu poprzez deaminację 5hmC jest mało prawdopodobny.

7. STRESZCZENIE

Deaminazy AID/APOBEC są rodziną białek mających unikatową zdolność wprowadzania mutacji zarówno do łańcuchów DNA jak i RNA. Obecnie najintensywniej badanym przedstawicielem tej rodziny jest deaminaza cytydyny indukowana aktywacją limfocytów B (ang. *activation-induced cytidine deaminase*, AID). Enzym ten, w aktywowanych antygenem limfocytach B, deaminując cytydynę (C) do urydyny, indukuje różnicowanie przeciwciał. Najnowsze dane literaturowe wskazują, że AID może również pełnić funkcje poza układem immunologicznym. Proponowany jest udział AID w procesach aktywnej demetylacji genomu albo poprzez deaminację 5-metylocytydyny (5mC) do tymidyny albo deaminację 5-hydroksymetylocytydyny (5hmC) do 5-hydroksymetylourydyny. W tym kontekście AID może mieć bezpośredni wpływ na proces reprogramowania epigenetycznego, regulację ekspresji genów i zjawisko pluripotencji komórek. Szereg doniesień potwierdza aktywność AID wobec modyfikowanych substratów. Jednocześnie, równie licznie, pojawiają się dane, które przeczą takiej możliwości. Głównym celem badań podjętych w ramach niniejszej pracy doktorskiej stała się zatem charakterystyka mechanizmów molekularnych leżących u podłoża aktywności biologicznej ludzkiej AID (hAID), a w szczególności jej aktywności enzymatycznej wobec C, 5mC i 5hmC *in vitro*.

W pierwszym etapie badań stworzono system produkcji hAID o pełnej aktywności enzymatycznej. Uzyskano w ten sposób zarówno hAID typu dzikiego jak i trzy jej mutanty. Następnie szczegółowo scharakteryzowano aktywności enzymatyczne tych białek wobec modyfikowanych i niemodyfikowanych substratów. Wyniki przeprowadzonych doświadczeń wykazały, że istnieją znaczące różnice w molekularnych mechanizmach rozpoznawania C i 5mC przez hAID. Zjawisko to było szczególnie wyraźne w przypadku mutantu N51A, który nie wykazywał aktywności wobec C ale był aktywny wobec 5mC. Jednocześnie wykazano brak aktywności badanych wariantów hAID wobec 5hmC. Przeprowadzone, w kolejnym etapie badań, analizy *in silico* struktury hAID oraz jej kompleksów z substratami pozwoliły zidentyfikować uwarunkowania strukturalne aktywności hAID wobec C, 5mC i 5hmC. Uzyskane wyniki dowodzą, że udział hAID w procesie aktywnej demetylacji, jak i innych procesach biologicznych związanych z deaminacją 5mC jest możliwy

ze względu na właściwości enzymatyczne tego białka. Jednocześnie wyniki te przeczą hipotezie o udziale hAID w procesach demetylacji DNA na drodze deaminacji 5hmC.

8. ABSTRACT

The AID/APOBEC deaminase family consists of enzymes that have unique ability to induce mutations in DNA and RNA chains. The activation-induced cytidine deaminase (AID) is currently the most extensively studied member of this family. In B cells activated by an antigen, AID initiates antibody diversification by deamination of cytidine (C) to uridine. The recently published data have suggested that AID might also play a role beyond the immune system. It has been proposed that AID initiates active genome demethylation by deamination of either 5-methylcytidine (5mC) to thymidine or 5-hydroxymethylcytidine (5hmC) to 5-hydroxymethyluridine. Consequently, AID might have a profound impact on epigenetic gene activation, regulation of gene expression and pluripotency. There are several lines of evidence that AID is active on modified substrates. Nevertheless, an almost equal number of contradictory observations exists. Therefore, the major aim of these studies was the characterization of molecular mechanisms underlying biological activity of human AID (hAID), especially its enzymatic activity on C, 5mC and 5hmC *in vitro*.

In the first stage of the studies, a method of production of a fully active hAID enzyme was developed. This procedure was applied for the production of wild type hAID and its three mutants. Subsequently, the enzymatic activities of these proteins on modified and non-modified substrates were assessed. The obtained results revealed that the structural determinants of C and 5mC recognition by hAID are different. It was particularly evident for the N51A mutant found to be inactive on C and active on 5mC. In addition, it was found that all examined hAID variants failed to deaminate 5hmC. The conducted *in silico* analysis of hAID structure and its complexes with the substrates allowed for the identification of the structural determinants of hAID activity on C, 5mC and 5hmC. The data presented here clearly showed that hAID has the properties that are expected for the enzyme that triggers active DNA demethylation by 5mC deamination. Moreover, 5mC may be targeted by hAID in other biological processes. At the same time, the obtained results provided another line of evidence that DNA demethylation by hAID-catalyzed 5hmC deamination is highly unlikely.

9. FINANSOWANIE



Program Międzynarodowe Projekty Doktoranckie „*Structural biology of plants and microbes*” realizowany dzięki wsparciu Unii Europejskiej, Europejskiego Funduszu Rozwoju Regionalnego, w ramach Programu Operacyjnego Innowacyjna Gospodarka.



Projekt „*Ochrona patentowa preparatu o aktywności enzymatycznej ludzkiej AID*” realizowany dzięki wsparciu Unii Europejskiej, Europejskiego Funduszu Rozwoju Regionalnego, w ramach Programu Operacyjnego Innowacyjna Gospodarka.



Krajowy Naukowy
Ośrodek Wiodący



Praca zrealizowana dzięki wsparciu Ministerstwa Nauki i Szkolnictwa Wyższego w ramach dotacji KNOW.

10. BIBLIOGRAFIA

1. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO & Honjo T (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem* **274**, 18470-18476.
2. Conticello SG (2012) Creative deaminases, self-inflicted damage, and genome evolution. *Ann N Y Acad Sci* **1267**, 79-85.
3. Samaranyake M, Bujnicki JM, Carpenter M & Bhagwat AS (2006) Evaluation of molecular models for the affinity maturation of antibodies: roles of cytosine deamination by AID and DNA repair. *Chem Rev* **106**, 700-719.
4. Conticello SG (2008) The AID/APOBEC family of nucleic acid mutators. *Genome Biol* **9**, 229.
5. Prochnow C, Bransteitter R, Klein MG, Goodman MF & Chen XS (2007) The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* **445**, 447-451.
6. Losey HC, Ruthenburg AJ & Verdine GL (2006) Crystal structure of *Staphylococcus aureus* tRNA adenosine deaminase TadA in complex with RNA. *Nat Struct Mol Biol* **13**, 153-159.
7. Rubio MA, Pastar I, Gaston KW, Ragone FL, Janzen CJ, Cross GA, Papavasiliou FN & Alfonzo JD (2007) An adenosine-to-inosine tRNA-editing enzyme that can perform C-to-U deamination of DNA. *Proc Natl Acad Sci U S A* **104**, 7821-7826.
8. Rogozin IB, Iyer LM, Liang L, Glazko GV, Liston VG, Pavlov YI, Aravind L & Pancer Z (2007) Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat Immunol* **8**, 647-656.
9. Harris RS, Petersen-Mahrt SK & Neuberger MS (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* **10**, 1247-1253.
10. Harris RS & Liddament MT (2004) Retroviral restriction by APOBEC proteins. *Nat Rev Immunol* **4**, 868-877.
11. Munk C, Willemsen A & Bravo IG An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. *BMC Evol Biol* **12**, 71.
12. Conticello SG, Thomas CJ, Petersen-Mahrt SK & Neuberger MS (2005) Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol* **22**, 367-377.
13. Sawyer SL, Emerman M & Malik HS (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol* **2**, E275.
14. Rosenberg BR, Hamilton CE, Mwangi MM, Dewell S & Papavasiliou FN (2011) Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. *Nat Struct Mol Biol* **18**, 230-236.
15. Smith HC, Bennett RP, Kizilyer A, McDougall WM & Prohaska KM (2011) Functions and regulation of the APOBEC family of proteins. *Semin Cell Dev Biol* **23**, 258-268.
16. Okuyama S, Marusawa H, Matsumoto T, Ueda Y, Matsumoto Y, Endo Y, Takai A & Chiba T Excessive activity of apolipoprotein B mRNA editing enzyme catalytic polypeptide 2 (APOBEC2) contributes to liver and lung tumorigenesis. *Int J Cancer* **130**, 1294-1301.
17. Prochnow C, Bransteitter R & Chen XS (2009) APOBEC deaminases-mutases with defensive roles for immunity. *Sci China C Life Sci* **52**, 893-902.
18. Rogozin IB, Basu MK, Jordan IK, Pavlov YI & Koonin EV (2005) APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. *Cell Cycle* **4**, 1281-1285.
19. Chaudhuri J & Alt FW (2004) Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol* **4**, 541-552.
20. Chiu YL & Greene WC (2008) The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol* **26**, 317-353.
21. Mbisa JL, Barr R, Thomas JA, Vandegraaff N, Dorweiler IJ, Svarovskaia ES, Brown WL, Mansky LM, Gorelick RJ, Harris RS, Engelman A & Pathak VK (2007) Human

- immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J Virol* **81**, 7099-7110.
22. Mariani R, Chen D, Schrofelbauer B, Navarro F, Konig R, Bollman B, Munk C, Nymark-McMahon H & Landau NR (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* **114**, 21-31.
23. Holmes RK, Malim MH & Bishop KN (2007) APOBEC-mediated viral restriction: not simply editing? *Trends Biochem Sci* **32**, 118-128.
24. Krupp A, McCarthy KR, Ooms M, Letko M, Morgan JS, Simon V & Johnson WE (2013) APOBEC3G polymorphism as a selective barrier to cross-species transmission and emergence of pathogenic SIV and AIDS in a primate host. *PLoS Pathog* **9**, e1003641.
25. MacDuff A & MacDuff R (2009) Pleural depth in medical patients: a radiological assessment. *Chest* **136**, 948-949.
26. Ikeda T, Abd El Galil KH, Tokunaga K, Maeda K, Sata T, Sakaguchi N, Heidmann T & Koito A Intrinsic restriction activity by apolipoprotein B mRNA editing enzyme APOBEC1 against the mobility of autonomous retrotransposons. *Nucleic Acids Res* **39**, 5538-5554.
27. Ikeda T, Ohsugi T, Kimura T, Matsushita S, Maeda Y, Harada S & Koito A (2008) The antiretroviral potency of APOBEC1 deaminase from small animal species. *Nucleic Acids Res* **36**, 6859-6871.
28. Vartanian JP, Henry M, Marchio A, Suspene R, Aynaud MM, Guetard D, Cervantes-Gonzalez M, Battiston C, Mazzaferro V, Pineau P, Dejean A & Wain-Hobson S Massive APOBEC3 editing of hepatitis B viral DNA in cirrhosis. *PLoS Pathog* **6**, e1000928.
29. Vieira VC & Soares MA The role of cytidine deaminases on innate immune responses against human viral infections. *Biomed Res Int* **2013**, 683095.
30. Holden LG, Prochnow C, Chang YP, Bransteitter R, Chelico L, Sen U, Stevens RC, Goodman MF & Chen XS (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121-124.
31. Bransteitter R, Prochnow C & Chen XS (2009) The current structural and functional understanding of APOBEC deaminases. *Cell Mol Life Sci* **66**, 3137-3147.
32. Bransteitter R, Pham P, Scharff MD & Goodman MF (2003) Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci U S A* **100**, 4102-4107.
33. McDougall WM & Smith HC Direct evidence that RNA inhibits APOBEC3G ssDNA cytidine deaminase activity. *Biochem Biophys Res Commun* **412**, 612-617.
34. Kohli RM, Abrams SR, Gajula KS, Maul RW, Gearhart PJ & Stivers JT (2009) A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J Biol Chem* **284**, 22898-22904.
35. Shivarov V, Shinkura R & Honjo T (2008) Dissociation of in vitro DNA deamination activity and physiological functions of AID mutants. *Proc Natl Acad Sci U S A* **105**, 15866-15871.
36. Endo Y, Marusawa H, Kinoshita K, Morisawa T, Sakurai T, Okazaki IM, Watashi K, Shimotohno K, Honjo T & Chiba T (2007) Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene* **26**, 5587-5595.
37. Chen KM, Harjes E, Gross PJ, Fahmy A, Lu Y, Shindo K, Harris RS & Matsuo H (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* **452**, 116-119.
38. Furukawa A, Nagata T, Matsugami A, Habu Y, Sugiyama R, Hayashi F, Kobayashi N, Yokoyama S, Takaku H & Katahira M (2009) Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J* **28**, 440-451.
39. Jaszczur M, Bertram JG, Pham P, Scharff MD & Goodman MF AID and Apobec3G haphazard deamination and mutational diversity. *Cell Mol Life Sci* **70**, 3089-3108.
40. Chelico L, Sacho EJ, Erie DA & Goodman MF (2008) A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. *J Biol Chem* **283**, 13780-13791.
41. Nowarski R, Britan-Rosich E, Shiloach T & Kotler M (2008) Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. *Nat Struct Mol Biol* **15**, 1059-1066.

42. Chelico L, Pham P, Calabrese P & Goodman MF (2006) APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA. *Nat Struct Mol Biol* **13**, 392-399.
43. Brar SS, Sacho EJ, Tessmer I, Croteau DL, Erie DA & Diaz M (2008) Activation-induced deaminase, AID, is catalytically active as a monomer on single-stranded DNA. *DNA Repair (Amst)* **7**, 77-87.
44. Bhagwat AS, Carpenter MA & Bujnicki JM (2008) Is AID a monomer in solution? *DNA Repair (Amst)* **7**, 349-350; author reply 351-342.
45. Sadofsky MJ (2001) The RAG proteins in V(D)J recombination: more than just a nuclease. *Nucleic Acids Res* **29**, 1399-1409.
46. Petersen-Mahrt SK, Harris RS & Neuberger MS (2002) AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99-103.
47. Di Noia J & Neuberger MS (2002) Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* **419**, 43-48.
48. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labeouze R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A & Durandy A (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**, 565-575.
49. Peled JU, Kuang FL, Iglesias-Ussel MD, Roa S, Kalis SL, Goodman MF & Scharff MD (2008) The biochemistry of somatic hypermutation. *Annu Rev Immunol* **26**, 481-511.
50. Krokan HE, Nilsen H, Skorpen F, Otterlei M & Slupphaug G (2000) Base excision repair of DNA in mammalian cells. *FEBS Lett* **476**, 73-77.
51. Budzko L JP, Figlerowicz M. (2013) Activation-induced cytidine deaminase (AID): single activity - pleiotropic effect. *BioTechnologia* **94**, 15-21.
52. Stavnezer J (2011) Complex regulation and function of activation-induced cytidine deaminase. *Trends Immunol* **32**, 194-201.
53. Mao Z, Bozzella M, Seluanov A & Gorbunova V (2008) DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902-2906.
54. de Yebenes VG & Ramiro AR (2006) Activation-induced deaminase: light and dark sides. *Trends Mol Med* **12**, 432-439.
55. Morgan HD, Dean W, Coker HA, Reik W & Petersen-Mahrt SK (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* **279**, 52353-52360.
56. Han JH, Akira S, Calame K, Beutler B, Selsing E & Imanishi-Kari T (2007) Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. *Immunity* **27**, 64-75.
57. Vuong BQ & Chaudhuri J Combinatorial mechanisms regulating AID-dependent DNA deamination: interacting proteins and post-translational modifications. *Semin Immunol* **24**, 264-272.
58. Sernandez IV, de Yebenes VG, Dorsett Y & Ramiro AR (2008) Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. *PLoS One* **3**, e3927.
59. Kinoshita K & Nonaka T (2006) The dark side of activation-induced cytidine deaminase: relationship with leukemia and beyond. *Int J Hematol* **83**, 201-207.
60. Yadav A, Olaru A, Saltis M, Setren A, Cerny J & Livak F (2006) Identification of a ubiquitously active promoter of the murine activation-induced cytidine deaminase (AICDA) gene. *Mol Immunol* **43**, 529-541.
61. Mai T, Zan H, Zhang J, Hawkins JS, Xu Z & Casali P (2010) Estrogen receptors bind to and activate the HOXC4/HoxC4 promoter to potentiate HoxC4-mediated activation-induced cytosine deaminase induction, immunoglobulin class switch DNA recombination, and somatic hypermutation. *J Biol Chem* **285**, 37797-37810.
62. Zan H & Casali P (2012) Regulation of Aicda expression and AID activity. *Autoimmunity* **46**, 83-101.

63. Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R, Fagarasan S, Honjo T & Nagaoka H (2009) B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. *Nat Immunol* **11**, 148-154.
64. Dedeoglu F, Horwitz B, Chaudhuri J, Alt FW & Geha RS (2004) Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. *Int Immunol* **16**, 395-404.
65. Lee-Theilen M & Chaudhuri J (2010) Walking the AID tightrope. *Nat Immunol* **11**, 107-109.
66. Pauklin S & Petersen-Mahrt SK (2009) Progesterone inhibits activation-induced deaminase by binding to the promoter. *J Immunol* **183**, 1238-1244.
67. Pauklin S, Sernandez IV, Bachmann G, Ramiro AR & Petersen-Mahrt SK (2009) Estrogen directly activates AID transcription and function. *J Exp Med* **206**, 99-111.
68. Maul RW & Gearhart PJ (2009) Women, autoimmunity, and cancer: a dangerous liaison between estrogen and activation-induced deaminase? *J Exp Med* **206**, 11-13.
69. Li G, Zan H, Xu Z & Casali P (2013) Epigenetics of the antibody response. *Trends Immunol* **34**, 460-470.
70. Teng G, Hakimpour P, Landgraf P, Rice A, Tuschl T, Casellas R & Papavasiliou FN (2008) MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* **28**, 621-629.
71. de Yebenes VG, Belver L, Pisano DG, Gonzalez S, Villasante A, Croce C, He L & Ramiro AR (2008) miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J Exp Med* **205**, 2199-2206.
72. Dorsett Y, McBride KM, Jankovic M, Gazumyan A, Thai TH, Robbiani DF, Di Virgilio M, Reina San-Martin B, Heidkamp G, Schwickert TA, Eisenreich T, Rajewsky K & Nussenzweig MC (2008) MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* **28**, 630-638.
73. Brar SS, Watson M & Diaz M (2004) Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J Biol Chem* **279**, 26395-26401.
74. Geisberger R, Rada C & Neuberger MS (2009) The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc Natl Acad Sci U S A* **106**, 6736-6741.
75. Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC & Reynaud CA (2008) Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* **205**, 1357-1368.
76. Orthwein A, Patenaude AM, Affar el B, Lamarre A, Young JC & Di Noia JM (2010) Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. *J Exp Med* **207**, 2751-2765.
77. McBride KM, Gazumyan A, Woo EM, Schwickert TA, Chait BT & Nussenzweig MC (2008) Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J Exp Med* **205**, 2585-2594.
78. Chaudhuri J, Khuong C & Alt FW (2004) Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* **430**, 992-998.
79. Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S, Li G, Schrum JP, Manis JP & Alt FW (2005) The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* **438**, 508-511.
80. McBride KM, Gazumyan A, Woo EM, Barreto VM, Robbiani DF, Chait BT & Nussenzweig MC (2006) Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proc Natl Acad Sci U S A* **103**, 8798-8803.
81. Cheng HL, Vuong BQ, Basu U, Franklin A, Schwer B, Astarita J, Phan RT, Datta A, Manis J, Alt FW & Chaudhuri J (2009) Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. *Proc Natl Acad Sci U S A* **106**, 2717-2722.
82. Longerich S, Tanaka A, Bozek G, Nicolae D & Storb U (2005) The very 5' end and the constant region of Ig genes are spared from somatic mutation because AID does not access these regions. *J Exp Med* **202**, 1443-1454.

83. Jaszczur M, Bertram JG, Pham P, Scharff MD & Goodman MF (2012) AID and Apobec3G haphazard deamination and mutational diversity. *Cell Mol Life Sci* **70**, 3089-3108.
84. Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, Thomas LM, Al-Qahtani A, White CA, Park SR, Steinacker P, Li Z, Yates J, 3rd, Herron B, Otto M, Zan H, Fu H & Casali P (2010) 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat Struct Mol Biol* **17**, 1124-1135.
85. Bachl J, Carlson C, Gray-Schopfer V, Dessing M & Olsson C (2001) Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J Immunol* **166**, 5051-5057.
86. Stavnezer J, Guikema JE & Schrader CE (2008) Mechanism and regulation of class switch recombination. *Annu Rev Immunol* **26**, 261-292.
87. Shen HM & Storb U (2004) Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled. *Proc Natl Acad Sci U S A* **101**, 12997-13002.
88. Yu K, Chedin F, Hsieh CL, Wilson TE & Lieber MR (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* **4**, 442-451.
89. Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR, Januszky K, Gregory RI, Deng H, Lima CD & Alt FW (2011) The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* **144**, 353-363.
90. El Hage A, French SL, Beyer AL & Tollervey D (2010) Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev* **24**, 1546-1558.
91. Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B, Barreto V, Nieland TJ, Root DE, Casellas R & Nussenzweig MC (2010) Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* **143**, 122-133.
92. Ronai D, Iglesias-Ussel MD, Fan M, Shulman MJ & Scharff MD (2005) Complex regulation of somatic hypermutation by cis-acting sequences in the endogenous IgH gene in hybridoma cells. *Proc Natl Acad Sci U S A* **102**, 11829-11834.
93. Michael N, Shen HM, Longerich S, Kim N, Longacre A & Storb U (2003) The E box motif CAGGTG enhances somatic hypermutation without enhancing transcription. *Immunity* **19**, 235-242.
94. Kothapalli N, Norton DD & Fugmann SD (2008) Cutting edge: a cis-acting DNA element targets AID-mediated sequence diversification to the chicken Ig light chain gene locus. *J Immunol* **180**, 2019-2023.
95. Keim C, Kazadi D, Rothschild G & Basu U (2013) Regulation of AID, the B-cell genome mutator. *Genes Dev* **27**, 1-17.
96. Nowak U, Matthews AJ, Zheng S & Chaudhuri J (2010) The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. *Nat Immunol* **12**, 160-166.
97. Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM & Nussenzweig MC (2003) C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell* **12**, 501-508.
98. Shinkura R, Ito S, Begum NA, Nagaoka H, Muramatsu M, Kinoshita K, Sakakibara Y, Hijikata H & Honjo T (2004) Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat Immunol* **5**, 707-712.
99. Durandy A, Peron S, Taubenheim N & Fischer A (2006) Activation-induced cytidine deaminase: structure-function relationship as based on the study of mutants. *Hum Mutat* **27**, 1185-1191.
100. Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, Imai K, Nonoyama S, Tashiro J, Ikegawa M, Ito S, Kinoshita K, Muramatsu M & Honjo T (2003) AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol* **4**, 843-848.
101. Wei M, Shinkura R, Doi Y, Maruya M, Fagarasan S & Honjo T (2011) Mice carrying a knock-in mutation of Aicda resulting in a defect in somatic hypermutation have impaired gut homeostasis and compromised mucosal defense. *Nat Immunol* **12**, 264-270.

102. Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA & Gehrke C (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* **10**, 2709-2721.
103. Yu DH, Ware C, Waterland RA, Zhang J, Chen MH, Gadkari M, Kunde-Ramamoorthy G, Nosavanh LM & Shen L (2013) Developmentally programmed 3' CpG island methylation confers tissue- and cell-type-specific transcriptional activation. *Mol Cell Biol* **33**, 1845-1858.
104. Robertson KD & Wolffe AP (2000) DNA methylation in health and disease. *Nat Rev Genet* **1**, 11-19.
105. Butler MG (2009) Genomic imprinting disorders in humans: a mini-review. *J Assist Reprod Genet* **26**, 477-486.
106. Ehrlich M (2002) DNA methylation in cancer: too much, but also too little. *Oncogene* **21**, 5400-5413.
107. Bestor TH (2000) The DNA methyltransferases of mammals. *Hum Mol Genet* **9**, 2395-2402.
108. Teperek-Tkacz M, Pasque V, Gentsch G & Ferguson-Smith AC (2011) Epigenetic reprogramming: is deamination key to active DNA demethylation? *Reproduction* **142**, 621-632.
109. Gehring M, Reik W & Henikoff S (2009) DNA demethylation by DNA repair. *Trends Genet* **25**, 82-90.
110. Fritz EL & Papavasiliou FN (2010) Cytidine deaminases: AIDing DNA demethylation? *Genes Dev* **24**, 2107-2114.
111. Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W & Walter J (2000) Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* **10**, 475-478.
112. Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J & Surani MA (2002) Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* **117**, 15-23.
113. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY & Blau HM (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042-1047.
114. Klug M, Heinz S, Gebhard C, Schwarzfischer L, Krause SW, Andreesen R & Rehli M (2010) Active DNA demethylation in human postmitotic cells correlates with activating histone modifications, but not transcription levels. *Genome Biol* **11**, R63.
115. Bruniquel D & Schwartz RH (2003) Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat Immunol* **4**, 235-240.
116. Zhu JK (2009) Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* **43**, 143-166.
117. Sanz LA, Kota SK & Feil R (2010) Genome-wide DNA demethylation in mammals. *Genome Biol* **11**, 110.
118. Zhu B, Zheng Y, Angliker H, Schwarz S, Thiry S, Siegmann M & Jost JP (2000) 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. *Nucleic Acids Res* **28**, 4157-4165.
119. Falnes PO, Johansen RF & Seeberg E (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* **419**, 178-182.
120. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L & Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935.
121. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, Abramowitz LK, Bartolomei MS, Rambow F, Bassi MR, Bruno T, Fanciulli M, Renner C, Klein-Szanto AJ, Matsumoto Y, Kobi D, Davidson I, Alberti C, Larue L & Bellacosa A (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67-79.
122. Guo JU, Su Y, Zhong C, Ming GL & Song H (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**, 423-434.
123. Nabel CS, Jia H, Ye Y, Shen L, Goldschmidt HL, Stivers JT, Zhang Y & Kohli RM (2012) AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol* **8**, 751-758.

124. Rangam G, Schmitz KM, Cobb AJ & Petersen-Mahrt SK (2012) AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud. *PLoS One* **7**, e43279.
125. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C & Xu GL (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303-1307.
126. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C & Zhang Y (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300-1303.
127. Bhutani N, Burns DM & Blau HM (2011) DNA demethylation dynamics. *Cell* **146**, 866-872.
128. Kriukiene E, Liutkeviciute Z & Klimasauskas S (2012) 5-Hydroxymethylcytosine--the elusive epigenetic mark in mammalian DNA. *Chem Soc Rev* **41**, 6916-6930.
129. Lin C, Yang L, Tanasa B, Hutt K, Ju BG, Ohgi K, Zhang J, Rose DW, Fu XD, Glass CK & Rosenfeld MG (2009) Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* **139**, 1069-1083.
130. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA & Cairns BR (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135**, 1201-1212.
131. Marr S, Morales H, Bottaro A, Cooper M, Flajnik M & Robert J (2007) Localization and differential expression of activation-induced cytidine deaminase in the amphibian *Xenopus* upon antigen stimulation and during early development. *J Immunol* **179**, 6783-6789.
132. Engel N, Tront JS, Erinle T, Nguyen N, Latham KE, Sapienza C, Hoffman B & Liebermann DA (2009) Conserved DNA methylation in *Gadd45a*(^{-/-}) mice. *Epigenetics* **4**, 98-99.
133. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE & Reik W (2010) Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101-1105.
134. Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, Honjo T, Nussenzweig A & Nussenzweig MC (2004) AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* **118**, 431-438.
135. Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S & Fischer RL (2007) DNA demethylation in the *Arabidopsis* genome. *Proc Natl Acad Sci U S A* **104**, 6752-6757.
136. Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP & Surani MA (2010) Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**, 78-82.
137. Larijani M, Frieder D, Sonbuchner TM, Bransteitter R, Goodman MF, Bouhassira EE, Scharff MD & Martin A (2005) Methylation protects cytidines from AID-mediated deamination. *Mol Immunol* **42**, 599-604.
138. Wijesinghe P & Bhagwat AS (2012) Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G. *Nucleic Acids Res* **40**, 9206-9217.
139. Abdouni H, King JJ, Suliman M, Quinlan M, Fifield H & Larijani M Zebrafish AID is capable of deaminating methylated deoxycytidines. *Nucleic Acids Res* **41**, 5457-5468.
140. Brenet F, Moh M, Funk P, Feierstein E, Viale AJ, Socci ND & Scandura JM (2011) DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS One* **6**, e14524.
141. Kumar R, DiMenna L, Schrode N, Liu TC, Franck P, Munoz-Descalzo S, Hadjantonakis AK, Zarrin AA, Chaudhuri J, Elemento O & Evans T (2013) AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* **500**, 89-92.
142. Morisawa T, Marusawa H, Ueda Y, Iwai A, Okazaki IM, Honjo T & Chiba T (2008) Organ-specific profiles of genetic changes in cancers caused by activation-induced cytidine deaminase expression. *Int J Cancer* **123**, 2735-2740.
143. Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, Azuma T, Okazaki IM, Honjo T & Chiba T (2007) *Helicobacter pylori* infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med* **13**, 470-476.
144. Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K & Honjo T (2003) Constitutive expression of AID leads to tumorigenesis. *J Exp Med* **197**, 1173-1181.

145. Yamanaka S, Balestra ME, Ferrell LD, Fan J, Arnold KS, Taylor S, Taylor JM & Innerarity TL (1995) Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc Natl Acad Sci U S A* **92**, 8483-8487.
146. Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH & Schatz DG (2008) Two levels of protection for the B cell genome during somatic hypermutation. *Nature* **451**, 841-845.
147. Yamane A, Resch W, Kuo N, Kuchen S, Li Z, Sun HW, Robbiani DF, McBride K, Nussenzweig MC & Casellas R (2011) Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol* **12**, 62-69.
148. Staszewski O, Baker RE, Ucher AJ, Martier R, Stavnezer J & Guikema JE (2011) Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig Loci in activated B cells. *Mol Cell* **41**, 232-242.
149. Chesi M, Robbiani DF, Sebag M, Chng WJ, Affer M, Tiedemann R, Valdez R, Palmer SE, Haas SS, Stewart AK, Fonseca R, Kremer R, Cattoretti G & Bergsagel PL (2008) AID-dependent activation of a MYC transgene induces multiple myeloma in a conditional mouse model of post-germinal center malignancies. *Cancer Cell* **13**, 167-180.
150. Klemm L, Duy C, Iacobucci I, Kuchen S, von Levezow G, Feldhahn N, Henke N, Li Z, Hoffmann TK, Kim YM, Hofmann WK, Jumaa H, Groffen J, Heisterkamp N, Martinelli G, Lieber MR, Casellas R & Muschen M (2009) The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia. *Cancer Cell* **16**, 232-245.
151. Kuppers R (2005) Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* **5**, 251-262.
152. Kuppers R & Dalla-Favera R (2001) Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* **20**, 5580-5594.
153. Ramiro AR, Jankovic M, Callen E, Difilippantonio S, Chen HT, McBride KM, Eisenreich TR, Chen J, Dickins RA, Lowe SW, Nussenzweig A & Nussenzweig MC (2006) Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. *Nature* **440**, 105-109.
154. Franco S, Gostissa M, Zha S, Lombard DB, Murphy MM, Zarrin AA, Yan C, Tepsuporn S, Morales JC, Adams MM, Lou Z, Bassing CH, Manis JP, Chen J, Carpenter PB & Alt FW (2006) H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. *Mol Cell* **21**, 201-214.
155. McCarthy H, Wierda WG, Barron LL, Cromwell CC, Wang J, Coombes KR, Rangel R, Elenitoba-Johnson KS, Keating MJ & Abruzzo LV (2003) High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. *Blood* **101**, 4903-4908.
156. Heintel D, Kroemer E, Kienle D, Schwarzinger I, Gleiss A, Schwarzmeier J, Marculescu R, Le T, Mannhalter C, Gaiger A, Stilgenbauer S, Dohner H, Fonatsch C & Jager U (2004) High expression of activation-induced cytidine deaminase (AID) mRNA is associated with unmutated IGVH gene status and unfavourable cytogenetic aberrations in patients with chronic lymphocytic leukaemia. *Leukemia* **18**, 756-762.
157. Shen HM, Peters A, Baron B, Zhu X & Storb U (1998) Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* **280**, 1750-1752.
158. Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, Chaganti RS, Klein U, Kuppers R, Rajewsky K & Dalla-Favera R (1998) BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci U S A* **95**, 11816-11821.
159. Koduru S, Wong E, Strowig T, Sundaram R, Zhang L, Strout MP, Flavell RA, Schatz DG, Dhodapkar KM & Dhodapkar MV Dendritic cell-mediated activation-induced cytidine deaminase (AID)-dependent induction of genomic instability in human myeloma. *Blood* **119**, 2302-2309.
160. Rai K, Sarkar S, Broadbent TJ, Voas M, Grossmann KF, Nadauld LD, Dehghanizadeh S, Hagos FT, Li Y, Toth RK, Chidester S, Bahr TM, Johnson WE, Sklow B, Burt R, Cairns BR & Jones DA (2010) DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. *Cell* **142**, 930-942.

161. Vecchio L, Seke Etet PF, Kipanyula MJ, Krampera M & Nwabo Kamdje AH Importance of epigenetic changes in cancer etiology, pathogenesis, clinical profiling, and treatment: what can be learned from hematologic malignancies? *Biochim Biophys Acta* **1836**, 90-104.
162. Feinberg AP & Vogelstein B (1983) Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun* **111**, 47-54.
163. Das PM & Singal R (2004) DNA methylation and cancer. *J Clin Oncol* **22**, 4632-4642.
164. Thompson MP & Kurzrock R (2004) Epstein-Barr virus and cancer. *Clin Cancer Res* **10**, 803-821.
165. Heath E, Begue-Pastor N, Chaganti S, Croom-Carter D, Shannon-Lowe C, Kube D, Feederle R, Delecluse HJ, Rickinson AB & Bell AI Epstein-Barr virus infection of naive B cells in vitro frequently selects clones with mutated immunoglobulin genotypes: implications for virus biology. *PLoS Pathog* **8**, e1002697.
166. Uchida J, Yasui T, Takaoka-Shichijo Y, Muraoka M, Kulwichit W, Raab-Traub N & Kikutani H (1999) Mimicry of CD40 signals by Epstein-Barr virus LMP1 in B lymphocyte responses. *Science* **286**, 300-303.
167. Li MJ & Maizels N (1999) Activation and targeting of immunoglobulin switch recombination by activities induced by EBV infection. *J Immunol* **163**, 6659-6664.
168. Kou T, Marusawa H, Kinoshita K, Endo Y, Okazaki IM, Ueda Y, Kodama Y, Haga H, Ikai I & Chiba T (2007) Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int J Cancer* **120**, 469-476.
169. Takai A, Toyoshima T, Uemura M, Kitawaki Y, Marusawa H, Hiai H, Yamada S, Okazaki IM, Honjo T, Chiba T & Kinoshita K (2009) A novel mouse model of hepatocarcinogenesis triggered by AID causing deleterious p53 mutations. *Oncogene* **28**, 469-478.
170. Khan M, Jahan S, Khaliq S, Ijaz B, Ahmad W, Samreen B & Hassan S Interaction of the hepatitis C virus (HCV) core with cellular genes in the development of HCV-induced steatosis. *Arch Virol* **155**, 1735-1753.
171. Chiba T & Marusawa H (2009) A novel mechanism for inflammation-associated carcinogenesis; an important role of activation-induced cytidine deaminase (AID) in mutation induction. *J Mol Med (Berl)* **87**, 1023-1027.
172. Marusawa H, Takai A & Chiba T Role of activation-induced cytidine deaminase in inflammation-associated cancer development. *Adv Immunol* **111**, 109-141.
173. Kim CJ, Song JH, Cho YG, Cao Z, Kim SY, Nam SW, Lee JY & Park WS (2007) Activation-induced cytidine deaminase expression in gastric cancer. *Tumour Biol* **28**, 333-339.
174. Machida K, Cheng KT, Pavio N, Sung VM & Lai MM (2005) Hepatitis C virus E2-CD81 interaction induces hypermutation of the immunoglobulin gene in B cells. *J Virol* **79**, 8079-8089.
175. Machida K, Kondo Y, Huang JY, Chen YC, Cheng KT, Keck Z, Fong S, Dubuisson J, Sung VM & Lai MM (2008) Hepatitis C virus (HCV)-induced immunoglobulin hypermutation reduces the affinity and neutralizing activities of antibodies against HCV envelope protein. *J Virol* **82**, 6711-6720.
176. Sohail A, Klapacz J, Samaranyake M, Ullah A & Bhagwat AS (2003) Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res* **31**, 2990-2994.
177. Abdouni H, King JJ, Suliman M, Quinlan M, Fifield H & Larijani M (2013) Zebrafish AID is capable of deaminating methylated deoxycytidines. *Nucleic Acids Res* **41**, 5457-5468.
178. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L & Schwede T (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252-258.
179. Kiefer F, Arnold K, Kunzli M, Bordoli L & Schwede T (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37**, D387-392.
180. Kurowski MA & Bujnicki JM (2003) GeneSilico protein structure prediction meta-server. *Nucleic Acids Res* **31**, 3305-3307.
181. Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815.

182. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L & Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802.
183. Foloppe N, MacKerell, Jr., Alexander D. (2010) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* **21**, 86-104.
184. Rappe AK, Casewit CJ, Colwell KS, Goddard WA & Skiff WM (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society* **114**, 10024-10035.
185. Cohen RM & Wolfenden R (1971) Cytidine deaminase from *Escherichia coli*. Purification, properties and inhibition by the potential transition state analog 3,4,5,6-tetrahydroimidine. *J Biol Chem* **246**, 7561-7565.
186. Chelico L, Pham P & Goodman MF (2009) Stochastic properties of processive cytidine DNA deaminases AID and APOBEC3G. *Philos Trans R Soc Lond B Biol Sci* **364**, 583-593.
187. Pham P, Bransteitter R, Petruska J & Goodman MF (2003) Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424**, 103-107.
188. Finney JL (2001) The water molecule and its interactions: the interaction between theory, modelling, and experiment. *Journal of Molecular Liquids* **90**, 303-312.
189. Tabbutt F (2001) Water: A Matrix of Life, 2nd Edition (Franks, Felix). *Journal of Chemical Education* **78**, 593.
190. http://www1.lsbu.ac.uk/water/water_molecule.html.
191. D. Xenides BRR, B.M. Rode (2006) Hydrogen bonding in liquid water: An ab initio QM/MM MD simulation study. *Journal of Molecular Liquids* **123**, 61-67.
192. Gajula KS, Huwe PJ, Mo CY, Crawford DJ, Stivers JT, Radhakrishnan R & Kohli RM (2014) High-throughput mutagenesis reveals functional determinants for DNA targeting by activation-induced deaminase. *Nucleic Acids Res* **42**, 9964-9975.
193. Dickerson SK, Market E, Besmer E & Papavasiliou FN (2003) AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* **197**, 1291-1296.
194. Papavasiliou FN & Schatz DG (2002) The activation-induced deaminase functions in a postcleavage step of the somatic hypermutation process. *J Exp Med* **195**, 1193-1198.
195. Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C & Neuberger MS (2004) Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. *J Mol Biol* **337**, 585-596.
196. Hu S, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, Shin J, Cox E, Rho HS, Woodard C, Xia S, Liu S, Lyu H, Ming GL, Wade H, Song H, Qian J & Zhu H (2013) DNA methylation presents distinct binding sites for human transcription factors. *Elife* **2**, e00726.

11. ZAŁĄCZNIKI

Załącznik 1.

Porównanie sekwencji dwudziestu klonów oligonukleotydu NNCN_2 (A, B), NNCN_3 (C, D) lub NNCN_4 (E, F) poddanego działaniu WT hAID (A, C, E) lub jej mutantu N51A (B, D, F). Znakiem „~” zaznaczono delecje / insercje wprowadzone przypuszczalnie przez zastosowaną podczas klonowania polimerazę; „~” oznaczono pozycje tożsame z sekwencją wyjściową; w miejscu zmiany widoczne są jednoliterowe symbole odpowiednie dla danego nukleotydu. Na rysunku nie przedstawiono sekwencji oskrzydających rozpoznawanych przez startery stosowane w reakcji PCR (zmiany w ich obrębie nie podlegają detekcji).

A.

NNCN_2 ATCAGTCGAA CCCTCGACCTG TCCTGCTCCATCCGCCCTACCGCCCATCGGTCATACGCTC A

klon 1~
klon 2~
klon 3~
klon 4 T. T. T. T. TT. T. T. T. T. T.
klon 5 .. T. TT. G. TT. TT. TT. TT. T. T. T. T. T.
klon 6 .. T. T. ... TTT. T. T. T. TT. T. TT. TT. TTT. T. T. T. T. T.
klon 7 .. T. G. T. T. T. T.
klon 8~
klon 9 .. T. T. ... TTT. T. T. T. T. T. A. T. T. T.
klon 10 .. T. T. ... T. T. T. T. T. T. T. T. T. T. T. T. T. T. T.
klon 11 .. T. T. TT. ATT. T. TT. ... TT. T. TT. TT. TT. TT. T. T. T. T. T. T. T. T.
klon 12 .. T. T. ... T. C. T. T. T. T. T. T. T. T. T. T. T. T. T. T. T. T.
klon 13~
klon 14 .. T. T. T. T. T. T.
klon 15~
klon 16~
klon 17 T. TT. T. T. T. T. T. T. T. T. T. T. G
klon 18 .. T. T. T. T. T. T. T. T. T. T. T. T. T. T. T.
klon 19~
klon 20~

B.

NNCN_2 ATCAGTCG AACCTCGACCTGTCCTG CTCTCCATCCGCCCTACCGCCCATCGGTCATACGCTCA

klon 1~
klon 2~
klon 3~
klon 4~
klon 5~
klon 6~
klon 7~
klon 8~
klon 9 G.
klon 10~
klon 11 .. G.~
klon 12~
klon 13~
klon 14~
klon 15~
klon 16~
klon 17~
klon 18~
klon 19~
klon 20~

C.

NNCN_3 AACAGGCGTCCACCCGAACTGCCATCCTCGCAACCAGGCTGGCATGCACCCTAACGTCCGCCCA
 klon 1
 klon 2
 klon 3 T. T. TT. T. TTT. T. T.
 klon 4 .. T. T. T. TTT. T. T. TT. T. T. TT. T. T. T. TTT. T. T. TT.
 klon 5 .. T. T. T. TTT. T. T. TT. T. T. TT. T. T. T. TTT. T. T. TT.
 klon 6 .. T. T. T. TTT. T. T. TT. T. T. TT. T. T. T. TTT. T. T. TT.
 klon 7 .. T. T. T. TT. T. T. TT. T. T. T. TTT. T. T. TT.
 klon 8
 klon 9
 klon 10 .. T. T. T. TTT. T. TT. TT. T. T. TT. T. T. T. TTT. T. TT. T. T.
 klon 11 .. T. T. T. TT. T. TT. TT. T. TT. T. TTT. T. T.
 klon 12 .. T. T. TT. TTT. T. TT. TT. T. T. TT. T. T. T. TTT. T. T. T.
 klon 13
 klon 14 TT. TTT. T. T. TTT. T. T.
 klon 15 T. TTT. T.
 klon 16
 klon 17 T. T. T. T. T. TT. T. T. TT. T. T. T. TTT. T. T. T. T.
 klon 18
 klon 19
 klon 20

D.

NNCN_3 AACAGGCGTCCACCCGAACTGCCATCCTCGCAACC-AGGCTGGCATGCACCCTAACGTCCGCCCA
 klon 1 C.
 klon 2
 klon 3 A. CG
 klon 4
 klon 5
 klon 6
 klon 7
 klon 8
 klon 9
 klon 10
 klon 11
 klon 12
 klon 13
 klon 14
 klon 15
 klon 16
 klon 17
 klon 18
 klon 19
 klon 20

E.

NNCN_4 ATCTGGCCTACTCCCCTTCTGCCGTGCCCGCGACCGCGCTTGC GCGCCTACAGTCTTCCCCTCT
 klon 1 ..T.....T.....T.....
 klon 2T.....
 klon 3
 klon 4T.....
 klon 5T.....
 klon 6T...T...TT...T...
 klon 7
 klon 8
 klon 9
 klon 10
 klon 11
 klon 12T.....T.....
 klon 13
 klon 14
 klon 15
 klon 16
 klon 17
 klon 18
 klon 19
 klon 20

F.

NNCN_4 ATCTGGCCTACTCCCCTTCTGCCGTGCCCGCGACCGCGCTTGC GCGCCTACAGTCTTCCCCTCT
 klon 1
 klon 2
 klon 3
 klon 4
 klon 5
 klon 6 ..A.....
 klon 7
 klon 8
 klon 9
 klon 10
 klon 11C.....
 klon 12
 klon 13
 klon 14
 klon 15
 klon 16
 klon 17
 klon 18
 klon 19
 klon 20

Załącznik 2.

Zestawienie oddziaływań pomiędzy atomami pierścienia pirymidynowego C lub 5mC a atomami reszt aminokwasowych tworzących kieszeń katalityczną WT hAID lub mutantu N51A.

reszta aminokwasowa	atom w obrębie danej reszty	atomy substratu w obrębie wygenerowanych kompleksów enzym:substrat (długość potencjalnego oddziaływania [Å] [#])			
		WT hAID:C	WThAID:5mC	N51A:C	N51A:5mC
H56	CA*	N3 (3,50) C2 (3,90) O2 (3,54)	O2 (3,52)	-	-
	CB	N3 (3,54)	-	N3 (3,52)	-
	CG	N3 (3,57) C2 (3,78) O2 (3,84)	O2 (3,96) C2 (3,83)	O2 (3,73) C2 (3,72) N3 (3,56)	N4 (3,89) C4 (3,77) N3 (3,47) C2 (3,87)
	CD	N3 (3,80) C2 (3,52) O2 (3,32)	C2 (3,49) O2 (3,40)	C2 (3,50) O2 (3,26)	C4 (3,88) N3 (3,48) C2 (3,50) O2 (3,61) N1 (3,93)
	N	-	O2 (3,94)	-	-
V57	HN	N3 (2,72) O2 (3,84)	O2 (3,86)	N3 (2,68) O2 (3,71)	N3 (1,91)
T27	OG	C6 (3,45) C5 (3,26) C4 (3,62)	C6 (3,71) C5 (3,32) C4 (3,59) C5M (3,49)	C5 (3,29) C4 (3,68)	C5 (3,58) C5M (3,69)
	CB	C5 (3,92) C4 (3,78)	C4 (3,89)	C4 (3,83)	C5 (3,97)
	CG	C4 (3,83) N4 (3,61)	N4 (3,69)	C4 (3,83) N4 (3,60)	C5 (3,46) C5M (3,66)
N51	CB	C2 (3,91)	C2 (3,93)	-	-
	ND	C2 (3,49) N1 (3,85)	C2 (3,40) N1 (3,75)	-	-
	2HD	O2 (2,20)	O2 (2,17)	-	-
A51	CB	-	-	C2 (3,77) O2 (3,33)	C2 (3,40) O2 (3,05) N1 (3,74)
W84	CB	-	C5M (3,95)	-	-
	CD1	-	-	-	C5M (3,62)
Y114	OH	-	-	C5 (3,92)	-

Podano wartości średnie z trajektorii MD. * Nomenklatura atomów jest zgodna z nomenklaturą stosowaną w programie Chimera 1.10.1, w którym wykonywane były analizy.

Załącznik 3.

Atomy znajdujące się w odległości van der Waalsa od atomów łańcucha bocznego reszty aminokwasowej N51 lub A51.

kompleks enzym:substrat	reszta N51 lub A51, atom tej reszty	reszta aminokwasowa w białku lub część substratu, atom	odległość [Å] [#]
WT hAID:AGCT	N51, CB*	V57, CG1	4,0
		T27, CB	4,2
		T27, CA	4,4
		T27, C	4,3
		T27, O	3,4
		T27, N	4,0
		E26, HA	4,4
		L52, N	3,2
		R50, C	3,8
		R50, O	4,3
		deoksycytydina, O2	3,1
		deoksycytydina, C2	4,0
		deoksycytydina, N1	4,3
		deoksycytydina, C1'	3,9
	deoksycytydina, O4'	3,9	
	N51, CG	T27, HB	4,0
		T27, CA	4,5
		T27, N	3,6
		T27, O	3,7
		E26, C	4,2
		E26, CA	4,0
		E26, HG2	4,2
		R50, C	4,4
		L52, N	3,9
		deoksycytydina, O2	4,4
		deoksycytydina, C1'	4,2
		deoksycytydina, O4'	3,8
	deoksycytydina, C4'	4,5	
	N51, OD1	T27, HB	4,3
		T27, CA	4,2
		T27, C	4,1
		T27, O	3,3
		T27, N	3,1
		E26, C	3,8
		E26, CA	3,5
		E26, CB	3,7
		E26, CG	4,1
		E26, OE2	4,5
		R50, C	4,3
		R50, HB2	3,8
		R50, HE	4,4
		deoksycytydina, O4'	4,5
	N51, ND2	T27, HN	4,0
E26, CA		4,3	
E26, HG2		4,3	
R25, O		4,5	
L52, N		4,0	

kompleks enzym:substrat	reszta N51 lub A51, atom tej reszty	reszta aminokwasowa w białku lub część substratu, atom	odległość [Å] [#]
WT hAID:AGCT	N51, ND2	deoksycytydina, C1'	4,1
		deoksycytydina, O4'	3,8
		deoksycytydina, C4'	4,0
		deoksycytydina, C3'	4,4
		deoksycytydina, O3'	3,7
		deoksytymidyna, O5'	4,0
		deoksytymidyna, C5'	3,7
		deoksytymidyna, O4'	4,4
	N51, HD21	T27, HN	4,1
		E26, CA	3,8
		E26, N	4,5
		E26, CG	4,5
		R25, C	4,4
		deoksycytydina, O4'	4,2
		deoksycytydina, C4'	4,1
		deoksycytydina, O3'	4,0
		deoksytymidyna, O5'	3,8
		deoksytymidyna, C5'	3,2
		deoksytymidyna, C4'	3,9
		deoksytymidyna, O4'	3,7
	N51, HD22	E26, HA	4,1
		L52, O	4,3
		L52, N	3,7
		deoksycytydina, C1'	3,7
		deoksycytydina, C2'	4,2
		deoksycytydina, C3'	4,0
		deoksycytydina, C4'	3,9
		deoksycytydina, O4'	3,7
		deoksycytydina, O3'	3,1
		deoksytymidyna, P	4,1
		deoksytymidyna, O5'	3,6
		deoksytymidyna, C5'	3,7
		WT hAID:AG5mCT	N51, CB
T27, CB	4,4		
T27, HG1	4,5		
T27, O	3,9		
T27, HN	3,8		
R50, C	3,7		
R50, O	4,3		
L52, N	3,3		
C55, O	4,3		
deoksycytydina, O2	3,2		
deoksycytydina, C2	3,9		
deoksycytydina, N1	4,2		
deoksycytydina, C1'	4,0		
deoksycytydina, O4'	3,9		
N51, CG	T27, CB		4,4
	T27, N		3,8
	T27, HG1		4,2
	T27, O		4,0
	T27, CA		4,5
	E26, HA		3,7
	R50, HB2		4,3
	R50, C		4,2

kompleks enzym:substrat	reszta N51 lub A51, atom tej reszty	reszta aminokwasowa w białku lub część substratu, atom	odległość [Å] [#]
WT hAID:AG5mCT	N51, CG	L52, N	4,0
		deoksycytydina, O2	4,2
		deoksycytydina, C1'	4,0
		deoksycytydina, O4'	3,5
		deoksycytydina, C4'	4,2
	N51, OD1	T27, CB	3,8
		T27, OG1	4,3
		T27, N	2,8
		T27, C	3,8
		T27, O	3,2
		T27, CA	3,6
		E26, C	3,7
		E26, CA	3,7
		E26, CB	4,3
		R50, HH21	4,5
		R50, CB	4,5
		R50, HA	4,3
		R50, C	4,1
		deoksycytydina, O4'	3,7
		deoksycytydina, C4'	4,3
	N51, ND2	T27, HN	3,9
		E26, HA	3,9
		L52, N	4,0
		deoksycytydina, C1'	4,1
		deoksycytydina, O4'	3,7
		deoksycytydina, C4'	4,0
	N51, HD21	deoksycytydina, O3'	4,0
		deoksytymidyna, C5'	4,2
		T27, HN	3,8
		E26, CA	4,3
		R25, O	4,3
		deoksycytydina, C1'	4,3
		deoksycytydina, O4'	3,7
		deoksycytydina, C4'	3,6
		deoksycytydina, C3'	4,3
		deoksycytydina, O3'	3,8
		deoksytymidyna, P	4,5
	deoksytymidyna, O2P	4,4	
	deoksytymidyna, O5'	4,5	
	deoksytymidyna, C5'	3,5	
	deoksytymidyna, C4'	4,3	
	N51, HD22	L52, O	4,1
		L52, C	4,2
		L52, CA	4,1
		L52, N	3,4
		deoksycytydina, O2	4,4
		deoksycytydina, C1'	4,1
deoksycytydina, O4'		4,0	
deoksycytydina, C4'		4,4	
deoksycytydina, O3'		3,8	
deoksytymidyna, P		4,5	
deoksytymidyna, O2P		4,0	
deoksytymidyna, C5'	4,4		
N51A:AGCT	A51, CB	V57, CG1	3,8

kompleks enzym:substrat	reszta N51 lub A51, atom tej reszty	reszta aminokwasowa w białku lub część substratu, atom	odległość [Å] [#]
N51A:AGCT	A51, CB	T27, CB	4,2
		T27, O	3,8
		T27, HN	3,7
		L60, HD11	4,4
		R50, C	3,7
		K52, N	3,3
		N53, HN	3,7
		deoksycytydyna, O2	3,0
		deoksycytydyna, C2	3,8
		deoksycytydyna, N1	4,4
		deoksycytydyna, C1'	4,3
		deoksycytydyna, O4'	4,1
N51A:AG5mCT	A51, CB	V57, CG1	3,6
		T27, CB	4,2
		T27, CA	4,4
		T27, N	3,9
		T27, O	3,7
		L60, HD12	4,0
		R50, C	3,7
		R50, O	4,3
		K52, N	3,4
		C55, O	4,2
		E26, HA	4,2
		E26, HG1	4,2
		deoksycytydyna, O2	3,2
		deoksycytydyna, C2	3,9
		deoksycytydyna, N1	4,4
		deoksycytydyna, C1'	4,2
deoksycytydyna, O4'	3,8		

Podano wartości średnie z trajektorii MD. * Nomenklatura atomów jest zgodna z nomenklaturą stosowaną w programie Chimera 1.10.1, w którym wykonywane były analizy.

Załącznik 4.

Atomy znajdujące się w odległości van der Waalsa od grupy metylowej (w przypadku 5mC) lub węgla C5 (w przypadku C).

kompleks enzym:substrat	reszta aminokwasowa w białku lub część substratu, atom	odległość od grupy metylowej [Å]	odległość od węgla C5 [Å] [#]
WT hAID:AGCT	H56, CD2*	—	4,1
	H56, CG		4,3
	T27, CG2		4,1
	T27, OG1		3,7
	T27, CB		4,2
	deoksycytyna, C1'		3,8
	deoksycytyna, C2'		4,3
	deoksycytyna, O4'		4,3
	atomy pierścienia cytozyny		<4,5
	WT hAID:AG5mCT		T27, CG2
T27, OG1		4,2	
W84, CB		4,0	
W84, CG		4,4	
W84, CD1		4,3	
S83, HD1		4,1	
deoksyguanozyna, O3'		3,9	
deoksyguanozyna, C2'		4,2	
deoksyguanozyna, C1'		4,3	
atomy pierścienia cytozyny		<4,5	
N51A:AGCT	H56, ND1	—	4,4
	H56, CG		4,5
	T27, CG2		4,0
	T27, OG1		3,7
	T27, CB		4,2
	deoksycytyna, C1'		3,7
	deoksycytyna, C2'		4,2
	deoksycytyna, O4'		4,2
	atomy pierścienia cytozyny		<4,5
	N51A:AG5mCT		T27, CG2
T27, OG1		3,9	
W84, CB		4,1	
deoksyguanozyna, O3'		4,4	
deoksyguanozyna, C2'		4,1	
deoksyguanozyna, C1'		4,0	
deoksyguanozyna, N3		4,3	
atomy pierścienia cytozyny		<4,5	

Podano wartości średnie z trajektorii MD. * Nomenklatura atomów jest zgodna z nomenklaturą stosowaną w programie Chimera 1.10.1, w którym wykonywane były analizy.

Załącznik CD.**Pliki video od 1 do 4**

Symulacje MD wygenerowanych modeli WT hAID (video 1 i 2) i mutantu N51A (video 3 i 4) z zadokowanym monobisfosforanem cytydyny (video 1 i 3) lub monobisfosforanem 5-metylocytydyny (video 2 i 4). Enzymy są reprezentowane przez szarą powierzchnię z widoczną strukturą drugorzędową. Monobisfosforany zadokowane do kieszeni katalitycznych są przedstawione w kolorze pomarańczowym. Różowy punkt reprezentuje kation cynkowy. Pokazano łańcuchy boczne wybranych reszt aminokwasowych (C87, C90, H56, E58, P86, Y114, W84, N51/A51) tworzących kieszeń katalityczną enzymów. W centrum katalitycznym widoczna jest również cząsteczka wody. Pliki video wygenerowano przy użyciu programu Chimera 1.10.1.

Pliki video od 5 do 8

Symulacje MD wygenerowanych modeli WT hAID (video 5 i 6) i mutantu N51A (video 7 i 8) z zadokowanym tetranukleotydem AGCT (video 5 i 7) lub AG5mCT (video 6 i 8). Widoczna jest struktura drugorzędowa białek (kolor niebieski). Pokazano łańcuchy boczne wybranych reszt aminokwasowych (C87, C90, H56, E58, Y114, W84, V57, N51/A51) tworzących kieszeń katalityczną enzymów. Resztę aminokwasową N51 lub A51 przedstawiono w kolorze białym, Y114 w kolorze błękitnym, W84 w kolorze ciemnoróżowym. Substrat zadokowany do kieszeni katalitycznej enzymu przedstawiono w następujący sposób: A – kolor zielony, G – kolor pomarańczowy, C lub 5mC – kolor żółty, T – kolor jasnoróżowy. Szary punkt reprezentuje kation cynkowy. W centrum katalitycznym widoczna jest również cząsteczka wody. Pliki video wygenerowano przy użyciu programu Chimera 1.10.1.