

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ВИСОКИХ ТЕХНОЛОГІЙ

Завідувач кафедри молекулярної біотехнології та біоінформатики
доцент, к.б.н. Нипорко Олексій Юрійович
протокол №__ засідання кафедри
від “__” _____ 2022 р.

**КЛЮЧОВІ МІШЕНІ МІКРОБІОМІВ ТА ВИВЧЕННЯ ЇХ ВПЛИВУ НА
ЦИТОСКЕЛЕТ ПІД ЧАС ПУХЛИНОУТВОРЕННЯ**

Випускна кваліфікаційна робота бакалавра
студентки спеціальності 091 «Біологія»
ОП «Біологія (високі технології)»

Коробко Поліни Сергіївни

Науковий керівник від кафедри:
доцент кафедри молекулярної
біотехнології та біоінформатики
к.б.н. **Самофалова Дарія Олексіївна**

Робота виконана під керівництвом наук. спів. відділу геноміки та молекулярної
біотехнології Державної Установи «Інститут харчової біотехнології та геноміки
НАН України», к.б.н Самофалової Д.О.

Оцінка захисту роботи

Київ – 2022 р.

АНОТАЦІЯ

Коробко П.С. Ключові мішені мікробіомів та вивчення їх впливу на цитоскелет під час пухлиноутворення. – Випускна кваліфікаційна робота бакалавра за спеціальністю 091 Біологія ОП «Біологія (високі технології)».

У роботі проведено аналіз структури та функцій мікробних білків sagA і fadA, проаналізовано інгібітори, проведено пошук гомологічних послідовностей у бактерій і у людини, проаналізовано сайти зв'язування і виконано молекулярний докінг. За умови підтвердження отриманих результатів, вони зможуть бути використані для розробки таргетної терапії пухлинних захворювань.

Ключові слова: sagA, fadA, мікробіомні мішені, таргетна терапія раку

ЗМІСТ

ЗМІСТ.....	3
ВСТУП.....	4
РОЗДІЛ 1 ОГЛЯД ЛІТЕРАТУРИ.....	6
1.1. Огляд захворюваності та смертності від раку і перспектив таргетної терапії.....	6
1.2. Загальна характеристика мікробіому людини.....	6
1.3. Вплив мікробіому на захворювання людини	8
1.4. Огляд протеїнів бактерій пов'язаних з раком	10
1.4.3. Протеїни бактерій, які впливають на цитоскелет під час онкогенезу.....	13
РОЗДІЛ 2 МАТЕРІАЛИ І МЕТОДИ.....	15
2.1. Загальний алгоритм дій	15
2.2. Структурне вирівнювання молекул <i>cagA</i> і <i>fadA</i> із молекулами інших бактерій та людини	15
2.3. Реконструкція 3D-моделей молекул <i>cagA</i> і <i>fadA</i>	15
2.4. Аналіз білок-білкових взаємодій	16
2.5. Проведення докінгу відібраних хітів.....	16
РОЗДІЛ 3.....	17
3.1. Аналіз структур <i>cagA</i> і <i>fadA</i> зі спорідненими послідовностями бактерій та людини та їх порівняння.....	17
3.2. Реконструкція та порівняння просторових структур.....	19
3.3. Аналіз білок-білкових взаємодій	20
3.4. Аналіз сайтів зв'язування на поверхні білків і проведення докінгу	21
ВИСНОВКИ	23
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	24

ВСТУП

Тема даного дослідження є актуальною на сьогодні, адже рак є однією з найрозповсюдженіших причин смертності у людини. Дослідження мішеней створює можливості для розробки таргетної терапії, що спрямована на мікробіомні білки, дає змогу більш ефективно та безпечніше для здорових клітин впливати на розвиток злоякісних пухлин.

Робота виконана в рамках наукового проекту Національного фонду досліджень під керівництвом наук. спів. відділу геноміки та молекулярної біотехнології Державної Установи «Інститут харчової біотехнології та геноміки НАН України».

Об'єктом дослідження є протеїни бактерій, що зв'язуються з протеїнами в організмі людини і викликають пошкодження цитоскелету клітин хазяїна такі як *fadA* адгезин і цитотоксично-асоційований імунодомінантний антиген *cagA*.

Метою роботи є дослідження мікробіомних мішеней, що є пов'язаними з із патогенетичними шляхами, які впливають на структуру цитоскелета, викликаючи запалення і розвиток ракових клітин. Знаходження інгібіторів до патогенних білків мікроорганізмів є перспективним для лікування та запобігання розвитку злоякісних пухлин.

Згідно з метою роботи були поставлені відповідні **завдання**, які полягають у:

1. Охарактеризувати мікробіомні мішені *cagA Helicobacter pylori* та *fadA Fusobacterium nucleatum*.
2. Виконати порівняння структур *cagA* та *fadA* зі спорідненими структурами інших бактерій, дослідити чи є гомологи даних молекул серед протеїнів *Homo sapiens*.
3. Виконати пошук просторових структур *cagA* та *fadA*.
4. Проаналізувати хімічні бази даних на наявність інгібіторів, дослідити їхню активність.

5. Виконати ліганд-білковий докінг та аналіз сайтів зв'язування.

РОЗДІЛ 1 ОГЛЯД ЛІТЕРАТУРИ

1.1. Огляд захворюваності та смертності від раку і перспективи таргетної терапії

Онкологічні захворювання значно впливають на здоров'я та тривалість життя людини. За статистикою захворюваності та смертності від раку на основі 185 країн і 36 видів раку, у 2020 році в усьому світі було зафіксовано 19,3 мільйона випадків раку та приблизно 10 мільйонів смертей спричинених раковими пухлинами. Захворюваність та смертність продовжують зростати зі старінням та збільшенням популяції [1].

Таргетна терапія пухлин є перспективним напрямком розвитку лікування, адже вона має переваги над традиційними методами терапії такими як хіміотерапія: не має токсичної дії на здорові клітини, не викликає резистентності. За допомогою таргетної терапії можливо: блокувати сигнали завдяки яким ростуть та розмножуються ракові клітини; руйнувати ракові клітини [2]. Насьогодні припускається, що приблизно 20% ракових захворювань людини можуть бути пов'язані з мікроорганізмами [3]. Тому дослідження протеїнів бактерій, які можуть запускати процес розвитку злоякісних пухлин, дає можливість для створення прицільно-діючої персоналізованої терапії.

1.2. Загальна характеристика мікробіому людини

Бактерії, гриби та віруси широко присутні на епітелії, який включає в себе: епідермальні покриви, ніздрі, бронхіальне дерево, протокову систему екзокринних органів таких як: молочні залози, піхва та шлунково-кишковий тракт. Бактерії регулярно перетинають слизову оболонку шлунково-кишкового

тракту і вступають в ентерогепатичну циркуляцію. Невелика кількість бактерій може циркулювати в крові час від часу, і деякі бактерії можуть концентруватися навіть в ракових пухлинах.

Мікробіом кишечника людини містить 10^{13} - 10^{14} організмів, а також унікальний геном, що містить 3×10^6 генів, що значно перевищує кількість генів людини. Бактеріальний мікробіом забезпечує різноманітність метаболічних функцій, які є важливими для організму хазяїна. Незважаючи на переважання мікробних клітин в людському тілі, вони мають невеликі розміри, і в сукупності вони складають лише кілька фунтів маси тіла кожної людини, що відповідає 2–7% біомаси людини без урахування ваги води.

У мікробіомі шлунково-кишкового тракту людини переважають чотири типи бактерій (рис. 1.1) : *Firmicutes* і *Bacteroidetes* (становлять 90%), *Proteobacteria* та *Actinomycetes* [4]. *Firmicutes* в основному грампозитивні, до них належать анаеробні клостридії, а також стрептококи та ентерококи. *Bacteroidetes* - це грамнегативні палички з гарною здатністю до перетравлювання складних вуглеводів, до яких належать *Bacteroides thetaiotamicron* та *Bacteroides fragilis*. Актинобактерії мають великий вміст ГЦ і є грампозитивними організмами, до яких належать біфідобактерії. *Proteobacteria* становлять різноманітну групу грамнегативних організмів і включають *Gammaproteobacteria*, *Escherichia coli*. [5]

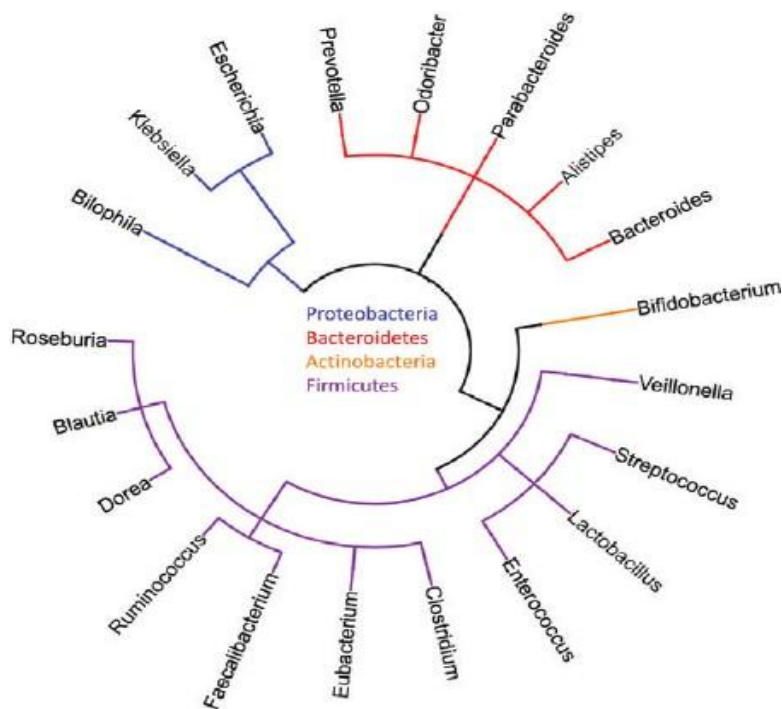


Рис. 1.1 – Основні типи і роди мікробіому кишечника людини [6]

1.3. Вплив мікробіому на захворювання людини

Взаємодія організму хазяїна і мікроорганізмів має велике значення для нормального фізіологічного розвитку та розвитку імунної системи і причетна до різноманітних захворювань, такі як запальні захворювання кишечника, рак, ожиріння та діабет 2 типу. Незважаючи на широкомасштабні дослідження типу «випадок-контроль», спрямовані на ідентифікацію мікробних таксонів або генів, що беруть участь у патогенезі, механізми, що пов'язують їх із захворюваннями, досі залишаються недослідженими [7].

Останнім часом завдяки розвитку техніки секвенування ДНК нового покоління було відкрито зв'язки між мікробіотою організму та різноманітними захворюваннями, такими як рак, серцево-судинні захворювання, хвороби, що пов'язані з обміном речовин (наприклад, ожиріння та цукровий діабет), розлади центральної нервової системи. Наприклад, мікроби кишечника відіграють роль

у формуванні та розвитку колоректального раку [8]. Значним фактором впливу мікробіому на захворювання є протеїн-протеїнові взаємодії між бактеріальними продуктами та білками хазяїна. Кілька канонічних білково-опосередкованих мікроб-асоційованих молекулярних шаблонів (МАМР), які безпосередньо запускають сигнальні шляхи організму хазяїна через рецептори розпізнавання патернів, що присутні на епітеліальних та імунних тканинах, зберігаються між патогенами та коменсалами, наприклад, між флагелліном та Toll-подібним рецептором 5 (TLR5). Припускається, що протеїн-протеїнові взаємодії з мікробними продуктами можуть бути поширеними і лежати в основі патогенезу мікробіомних розладів [7].

13,5% бактеріальних кластерів, пов'язаних із захворюванням, не містять членів з анотованими мікробними шляхами/функціями в KEGG (Кіотська енциклопедія генів і геномів). Використовуючи пошук гомології щодо бактеріальних взаємодій, більшість із цих генів можна анотувати відповідно до шляхів їхньої людської мішені, отримуючи ймовірний молекулярний механізм, пов'язаний із захворюванням. Бактеріальні кластери, що є пов'язаними із захворюваннями, зазвичай мають ортологи білків у людини, які виконують вторинні функції у доповнення до основних у клітині. Зазвичай ці кластери беруть участь у трансляції та центральному метаболізмі. Наприклад, *Mycoplasma pneumoniae* GroEL і *Streptococcus suis* енолаза (білок, який бере участь у гліколізі) зв'язуються як з плазміненом людини, так і з компонентами позаклітинного матриксу. *Mycobacterium tuberculosis* DnaK сигналізує лейкоцитам, викликаючи вивільнення хемокінів CCL3-5. Гліцеральдегід-3-фосфатдегідрогеназа *Streptococcus pyogenes* (GAPDH), канонічно бере участь у гліколізі, і може бути переміщена на поверхню клітини, де вона відіграє роль адгезину, а також може сприяти апоптозу клітини людини [9]. Ці приклади чітко ілюструють, як бактеріальні «білки домашнього господарства» використовуються патогенами для впливу на здоров'я людини.

Багато білків людини, пов'язаних із захворюваннями, є відомими мішенями для ліків. Наприклад, імаїнібу мезилат діє на декілька тирозинкіназ

сімейства Src, включаючи LCK, який бере участь у розвитку Т-клітин і має визнану роль у запаленні [10]. Бактеріальні білки, спрямовані на ці самі кінази, постійно присутні в контрольних групах як при запальному захворюванні кишечника, так і при колоректальному раку [7]. Крім того, іматиніб також може зупинити проліферацію пухлинних клітин товстої кишки і, як правило, бере участь у запальних шляхах через інгібування продукції TNF-альфа [11]. У відповідності з ідеєю, що бактеріальні білки можуть діяти терапевтично, є прямі докази принаймні двох комменсальних білків, які викликають фізіологічний вплив на господаря при введенні через зонд: очищений *A. muciniphila* Amuc_1100 і *F. prausnitzii* MAM для зменшення непереносимості глюкози і зменшення коліту [12]. Використання як ліків метаболітів, отриманих з мікробіому, є добре оціненою ідеєю [13], і дослідження білкових продуктів, що діють через протеїн-протеїнові взаємодії дає можливість до розроблення нових таргетно-спрямованих ліків.

1.4. Огляд протеїнів бактерій пов'язаних з раком

Деякі види бактерій мікробіому людини мають здатність виділяти патогенні продукти. Припускається, що ці продукти виділення причетні до апоптозу клітин, генотоксичності, ухилення від імунного захисту, запальних процесів і, зрештою, розвитку раку [13]. Патогенні бактерії можуть взаємодіяти з людським організмом, викликаючи хронічне запалення та подальше пошкодження, яке було пов'язано з онкогенезом.

Специфічні фактори вірулентності були ідентифіковані як потенційні локуси генів, важливі для стимулювання канцерогенної активності різних злоякісних пухлин. Сюди входять такі, які виробляються видами бактерій, такими як *H. pylori*, *F. nucleatum* і *B. fragilis*, а також *P. gingivalis*, що викликають рак шлунку, колоректальний рак і рак підшлункової залози.

1.4.2. Протеїни, що виділяє *Helicobacter pylori*

H. pylori є встановленим фактором розвитку злоякісної пухлини шлунка, і його канцерогенна активність визначається кількома факторами вірулентності. Найбільш вивчені на сьогоднішній день: *cagA* (цитотоксичний ген А), *vacA* (вакуолізуючий цитотоксин А) та *OipA* (зовнішній запальний білок) [14]. *CagA* функціонує шляхом ін'єкції в епітеліальні клітини шлунка фактором секреції IV типу (Т4SS) і одноразово впливає на внутрішньоклітинні системи трансдукції, що дозволяє реорганізувати актиновий цитоскелет [15]. *CagA* також стимулює вивільнення прозапального цитокіну ІЛ-8, що призводить до подальших змін в епітелії шлунка та можливої ініціації надмірного розростання клітин. Таке тривале запалення може призвести до хронічного гастриту та пухлиноутворення [16]. *VacA*, також важливий токсин, може сприяти колонізації *H. pylori* в клітинах шлунка і пов'язаний з прогресуванням захворювання. Припускається, що *VacA* збільшує активність уреазу за рахунок своєї здатності створювати клітинні вакуолі, що сприяє дифузії сечовини із тканини слизової оболонки в просвіт шлунка. Це забезпечує виживання *H. pylori* у кислому середовищі шлунка та запобігає клітинному апоптозу. Білок *OipA*, що кодується геном *oipA*, також може брати участь у стимуляції ІЛ-8 та хронічному запаленні і може працювати синергетично з *CagA* у розвитку раку шлунку [17].

1.4.1. Протеїни, що виділяє *Fusobacterium nucleatum*

Fusobacterium nucleatum присутній у мікробіомі ротової порожнини і є патогеном ясен. Ці бактерії пов'язані з пародонтитом, апендицитом, а також із колоректальним раком і були ідентифіковані як нова потенційна терапевтична мішень і прогностичний біомаркер для цього виду раку.

Білки клітинної поверхні *F. nucleatum*, FadA (Fusobacterium adhesion A), Far2 (білок активації фібробластів 2) і radD (ген радіації) є факторами вірулентності, які беруть участь у розвитку колоректального раку. Дисперсія адгезійного комплексу FadA за межі слизової оболонки порожнини рота пов'язана із запаленням та проліферацією клітин. Far2 і RadD, члени сімейства аутоотранспортерних білків, також відіграють важливу роль у вірулентності та адгезивній природі *F. Nucleatum* [18]. Far2 ідентифікує та приєднується до пухлин, що демонструють Gal-GalNAc (ацетилгалактозамін), який надмірно експресується як при колоректальній аденокарциномі, так і при метастазах КРР [19]. Far2 має захисний механізм через інактивацію природних клітин-кілерів проти злоякісних клітин, викликану пригніченням рецептора TIGIT [19]. RadD пов'язаний із загибеллю клітин і агрегацією клітина-клітина.

1.4.2. Протеїни, що виділяє *Escherichia coli*

Escherichia coli знаходиться в мікробіомі кишечника людини і може бути як нешкідливим коменсалом, так і потенційно небезпечним патогеном. Цикломодуліни, що утворюються патогенною *E. coli*, включають колібактин, CDT (цитолетальний розтягуючий токсин) і CNF (цитотоксичний некротизуючий фактор). Цикломодуліни є генотоксичними, вони порушують регуляцію клітинного циклу і беруть участь як у клітинній диференціації, так і в апоптозі, що є допоміжними механізмами в онкогенезі товстої кишки [20]. Колібактин – це сполука, синтезована ферментами, кодованими локусом Pks, що містить до 18 генів, включаючи колібактин А (clbA) і колібактин Р (clbP), які, ймовірно, відповідають за вірулентність. *E. coli* відіграє роль у розвитку колоректального раку, зокрема штамми, що продукують колібактин, що було підтверджено як у дослідженнях *in vitro*, так і *in vivo* на тваринах [21].

Дослідження показали, що 86% штамів *E.coli* (фенотипу B2), виділених із тканини аденокарциноми товстої кишки 48 людей, були цикломодулін-

позитивними [91]. І CDT, і колібактин мають ДНКазну активність, створюючи розриви дволанцюгової ДНК в епітеліальних клітинах [18]. Також встановлено, що колібактин, що виділяє *E. coli*, сприяє проліферації клітин і дозріванню пухлини під час колоректального раку. Припускається, що цей процес змінює мікросередовище пухлини і що фактори росту (HGF (фактор росту гепатоцитів), FGF (фактор росту фібробластів) і GM-CSF (гранулоцит-макрофаг-колонієстимулюючий фактор)) зі старіючих клітин можуть бути активовані, для стимуляції розвитку пухлини [22].

Функція CNF пов'язана із ненормальним зростанням і диференціюванням клітин. CNF активує перехід клітинного циклу G1-S і таким чином підвищує реплікацію ДНК. Загальна кількість клітин не збільшується, а навпаки, клітини розвиваються в багатоядерний формат, таким чином, цей токсин може перешкоджати диференціації клітин і згодом викликати загибель клітини [23].

1.4.3. Протеїни бактерій, які впливають на цитоскелет під час онкогенезу

До протеїнів, що виділяються бактеріями і впливають на цитоскелет належать *fadA* і *cagA*. *fadA* впливаючи на протеїн людини E-кадгерин активує бета-катенінові сигнальні шляхи.

Адгезійний комплекс кадгерин-катенін є ключовим компонентом міжклітинного спайкового з'єднання, яке сприяє як стабільності тканин, так і динамічним рухам клітин в епітеліальних і неепітеліальних тканинах. Адгезійний комплекс кадгерину з'єднує сусідні клітини та актин-міозиновий цитоскелет і, таким чином, сприяє механічному з'єднанню між клітинами, що викликає багато морфогенетичних подій і відновлення тканин. Механотрансдукція при спайках кадгерину дозволяє клітинам відчувати, сигналізувати та реагувати на фізичні зміни в їхньому середовищі. Центральне місце в цьому процесі займає динамічний зв'язок комплексу з актиновими

філаментами (F-актин), які самі по собі є структурно динамічними і піддаються напрузі. FadA модулює E-кадгерин і активує передачу сигналів β -катеніну, що призводить до посилення експресії факторів транскрипції, онкогенів, генів запалення, а також до стимуляції росту клітин колоректального раку.

Адгезин FadA у *F. nucleatum* і CagA у *H. pylori* асоціюються з E-кадгерином, що призводить до порушення регуляції сигнального шляху β -катеніну, викликаючи ядерну транслокацію β -катеніну, таким чином активуючи гени-мішені, що призводить до онкогенезу. При порушенні цього сигнального шляху відбувається фізична асоціація патогенних молекул клітинної адгезії з різними компонентами цитоскелета організму-хазяїна. Це сприяє поширенню запалення і пригніченню імунітету інфікованого організму [24].

РОЗДІЛ 2

МАТЕРІАЛИ І МЕТОДИ

2.1. Загальний алгоритм дій

Відбір структур для подальшого аналізу. Реконструкція 3D-моделей цільових білків. Оцінка просторових структур обраних мішеней і їх якості. Пошук інгібіторів та їх аналіз. Аналіз сайтів зв'язування та проведення молекулярного докінгу.

2.2. Структурне вирівнювання молекул *cagA* і *fadA* із молекулами інших бактерій та людини

Амінокислотні послідовності білків *cagA* і *fadA* були відібрані з бази даних UniProt (<https://www.uniprot.org/>). Пошук гомологічних послідовностей у інших видів бактерій та у людини було виконано за допомогою інструменту пошуку схожих послідовностей BLAST (Basic Local Alignment Search Tool) (<https://www.uniprot.org/blast/>). Вирівнювання послідовностей було виконано в програмі MEGA (<https://www.megasoftware.net/>) із використанням алгоритму ClustalW (<http://www.ebi.ac.uk/clustalw/>).

2.3. Реконструкція 3D-моделей молекул *cagA* і *fadA*

3D-моделі молекул *cagA* і *fadA* були отримані із сайту банку 3D-структур білків та нуклеїнових кислот RCSB PDB (<https://www.rcsb.org/>) і проаналізовані. Для отримання кращої якості структур були реконструйовані *de novo* на основі послідовностей із Uniprot білки-мішені. Реконструкція здійснювалась у сервісі SwissModel (<https://swissmodel.expasy.org/>). Валідація отриманої структури була

здійснена за допомогою MolProbity (<http://molprobity.biochem.duke.edu/>). При цьому оцінювались такі показники як 'all atoms Clashscore', 'Poor rotamers', 'Ramachandran outliers'. Подібність реконструйованих структур та структур з Protein Data Bank була оцінена за допомогою інструменту LDDT (Local Distance Difference Test) у Swiss-model.

2.4. Аналіз білок-білкових взаємодій

Наявні білок-білкові взаємодії шукали за допомогою сервісів Uniprot та NCBI (<https://www.ncbi.nlm.nih.gov/>). Для візуалізації білок-білкових взаємодій було використано сервіс IntAct (<https://www.ebi.ac.uk/intact/home>).

2.5. Проведення докінгу відібраних хітів

Докінг був зроблений у програмі Schrödinger (ліцензія НВП “І.Ф.ЛАБ”) за допомогою модуля Glide [27]. Мінімізація енергії відібраних хітів проводилась за допомогою силового поля OPLS3e (Optimized Potentials for Liquid Simulations) в модулі Ligprep. Карта сайтів зв’язування була побудована за допомогою Receptor grid generation у плагіну Glide Maestro (ліцензія НВП “І.Ф.ЛАБ”).

РОЗДІЛ 3

РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ ТА ЇХ ОБГОВОРЕННЯ

3.1. Аналіз структур *cagA* і *fadA* зі спорідненими послідовностями бактерій та людини та їх порівняння

За допомогою пошуку в BLAST було виявлено споріднені послідовності людини і бактерій до *fadA* і *cagA*. В результаті пошуку споріднених до *fadA* амінокислотних послідовностей було знайдено 250 бактеріальних послідовностей і з них було відібрано 10 послідовностей з такими показниками ідентичності: адгезин *fadA Fusobacterium hwasookii* (99,2%), *fadA Fusobacterium periodonticum* (94.6%), *fadA Fusobacterium canifelinum* 98.4%, білок рекомбінації ДНК RmuC *Planctomycetes bacterium* (30,8%), *fadA Leptotrichia hofstadii* (35.7%), *fadA Pseudoleptotrichia goodfellowii* (29.5%), *fadA Leptotrichia wadei* (29.9%), *fadA Leptotrichia sp. oral taxon* (34.1%), пептидоглікан-зв'язуючий білок *Alphaproteobacteria bacterium* (32.4%). Результати вирівнювання підтвердили результати аналізу у BLAST і можна припустити консервативність послідовностей *fadA F. nucleatum*.

Під час пошуку споріднених амінокислотних послідовностей до *fadA* було відібрано 1 білок у людини з показником ідентичності 33,7% міозин-9 *Homo sapiens*.

При пошуку споріднених послідовностей бактерій до *cagA* було відібрано 2 високоідентичні структури різних штамів *Helicobacter pylori* з такими показниками ідентичності: *cagA Campylobacter pylori* (90.0%), *cagA* штаму J99 (85.8%).

При пошуку послідовностей людини споріднених до *cagA* було знайдено лише молекули з низьким ступенем ідентичності (<30%), тому для вирівнювання вони не використовувались.

Результати множинного вирівнювання (Рисунки 3.1, 3.2) дозволяють припустити, що молекули *fadA* і *cagA* мають консервативні каталітичні домени.

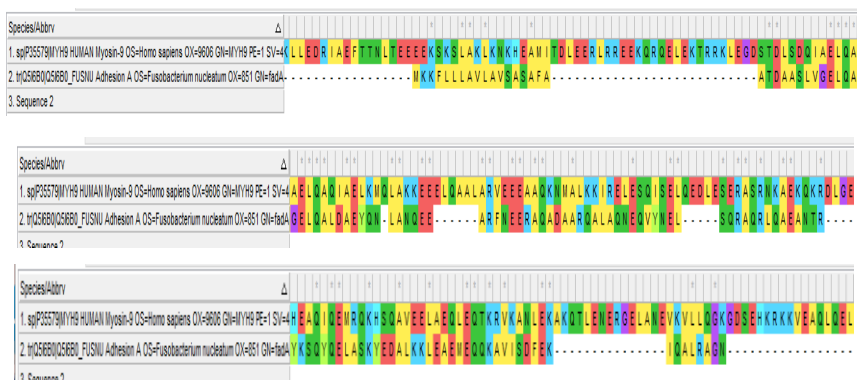


Рис. 3.1 – Множинне вирівнювання послідовностей людини споріднених з *fadA*



Рис 3.2 – Множинне вирівнювання послідовностей бактерій споріднених з *fadA*.

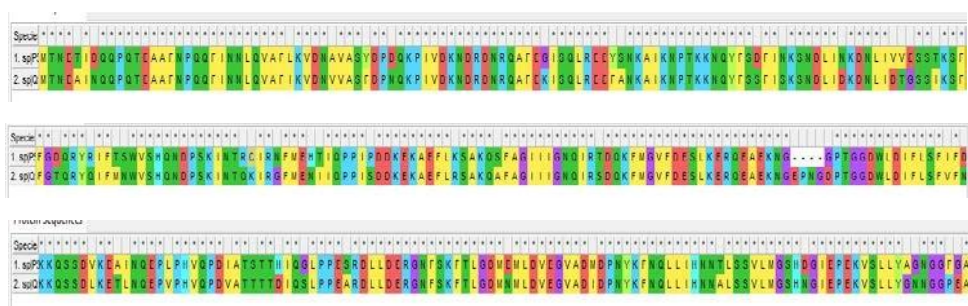


Рис. 3.2 – Множинне вирівнювання послідовностей бактерій споріднених з *cagA*.

3.2. Реконструкція та порівняння просторових структур

Просторові структури *cagA* і *fadA* були відібрані з сайту банку даних трьохвимірних структур білків RCSB PDB. Інформація про показники структури взята зі звіту валідації структурних моделей в PDB.

Якість структури *cagA* становить 3.19 Å. Clashscore для всіх атомів становить 2, що говорить про гарну якість відібраної структури. На карті Рамачандрана, де було проаналізовано 441 із загальної кількості 569 залишків (78%) [25] зазначено такі показники структури: 95% залишків знаходяться у бажаних областях (417/441), 5% у дозволених областях (21/441).

Якість структури *fadA* становить 2.00 Å. Clashscore всіх атомів становить 3, отже структура є якісною. На карті Рамачандрана [26] було проаналізовано 107 із 119 залишків (90%), з яких у бажаних областях знаходяться 99% (106/107), а у дозволених 1% (1/107).

Для подальшого використання структури білків *cagA* і *fadA* були реконструйовані *de novo* і проведене порівняння реконструйованих структур із готовими моделями взятими з PDB. Щоб побудувати послідовність *fadA F. nucleatum* була використана послідовність Q5I6B0 з бази даних білкових послідовностей UniProt. Реконструкція моделі була виконана на веб-сервері SWISS-MODEL. Реконструйований білок складається зі 129 амінокислотних залишків. Якість структури становить 2.00 Å. Clashscore становить 0 що є кращим результатом у порівнянні зі структурою з PDB. Показник QMEANDisCO Global становить 0.82. На карті Рамачандрана у бажаних областях знаходяться 92.37% залишків,

Для побудови послідовності *cagA H. pylori* була використана послідовність P55980 з бази даних UniProt. Реконструкція моделі була виконана на веб-сервері SWISS-MODEL. Реконструйований білок складається з 1186 амінокислотних залишків. Якість структури становить 3.30 Å. Clashscore становить 0.55, що є кращим результатом у порівнянні зі структурою з PDB.

Показник QMEANDisCO Global становить 0.71. На карті Рамачандрана у бажаних областях знаходяться 99.06% залишків.

Подібність реконструйованих структур та структур з PDB оцінювалась за допомогою Local Distance Difference Test. Показник Global IDDT становив для структур *cagA* 0.9998, а для *fadA* 0.6354

3.3. Аналіз білок-білкових взаємодій

Аналіз білок-білкових взаємодій був проведений за допомогою веб-сервісу IntAct.

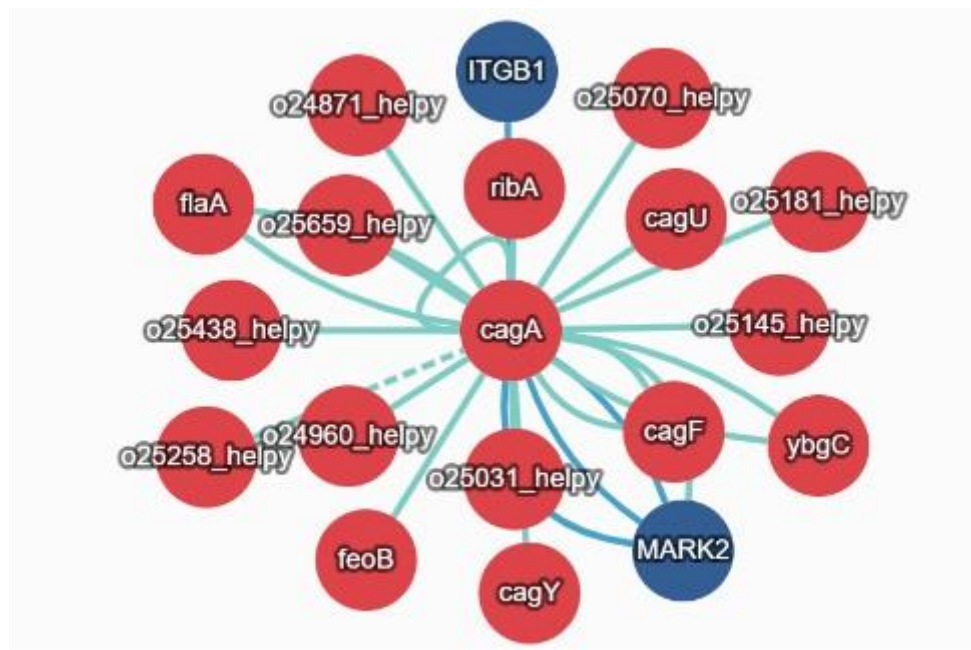


Рис. 3.3 – Візуалізація білок-білкових взаємодій для *cagA*

На рисунку 3.3. показано взаємодію *cagA* з іншими білками, зокрема з білками людини серин-треонін протеїнкіназою MARK2 і 1- β -інтегрином. MARK2 бере участь у регуляції полярності клітин і динаміки мікротрубочок. 1- β -інтегрин зв'язується з інтегрином- α -1 та інтегрином- α -2, утворюючи комплекси інтегрину, які функціонують як рецептори колагену.

Для *fadA* було знайдено білок-білкові взаємодії між таким білком людини як E-кадгерин, який бере участь у механізмах регулювання міжклітинної адгезії, клітинної рухливості та проліферації епітеліальних клітин.

3.4. Аналіз сайтів зв'язування на поверхні білків і проведення докінгу

Для молекули *sagA* було проаналізовано 5 сайтів зв'язування, для яких показник SiteScore становив: 0.988, 0.988, 0.993, 1.036, 1.045. Dscore становив: 1.013, 1.011, 1.021, 0.997, 1.125. Volume становив 235.984, 229.467, 476.427, 772.093, 477.799. З усіх сайтів зв'язування проявив активність лише п'ятий, що корелює з попередньою оцінкою. Сайти зв'язування знаходились біля таких положень: *Pe 51, Ala 168, Gln 410, Leu 727, Leu 380*.

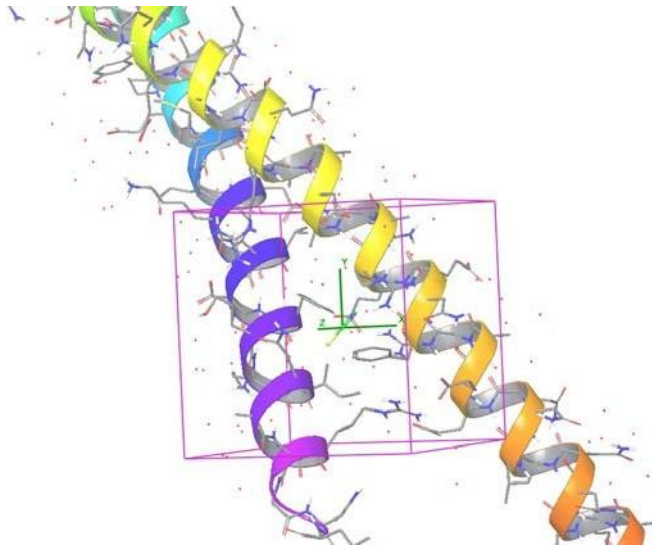


Рис. 3.4 – Дослідження сайту взаємодії з лігандами для структури *fadA*

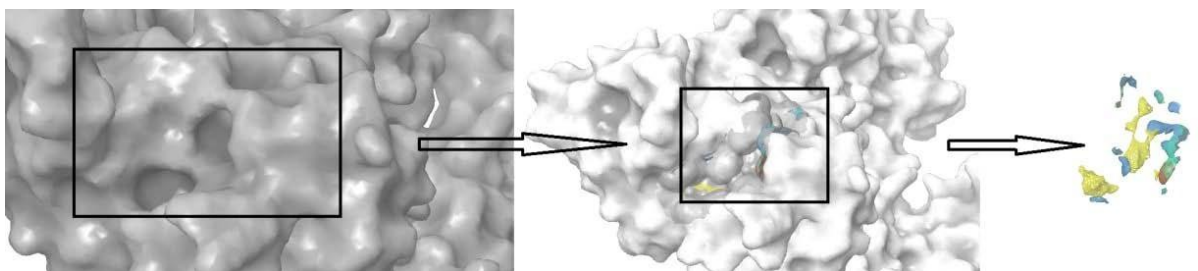


Рис. 3.5 – Сайт зв'язування білку *sagA*

Для молекули *fadA* було знайдено 1 потенційний сайт зв'язування. Для білку *fadA* дослідження єдиного потенційного сайту взаємодії не дало результату, тому на основі відомих сполук і встановлених даних було розроблено бібліотеку аналогів. Отримані результати можуть в подальшому бути використані для розробки таргетної терапії раку.

ВИСНОВКИ

Досліджено, що білки *cagA* та *fadA* впливають на цитоскелет шляхом асоціації з E-кадгерином, що порушує регуляцію β -катенінового сигнального шляху і призводить до пухлиноутворення.

1. Було проведене дослідження споріднених структур до *cagA* і *fadA*. На основі проведеного множинного вирівнювання було встановлено, що і сполука *cagA*, і сполука *fadA* є консервативними по відношенню до споріднених послідовностей людини та інших видів бактерій.

2. Було проаналізовано якість та порівняно 3D-структури білків реконструйовані *de novo* зі структурами встановленими експериментально і показано, що реконструйовані структури мають кращі якісні показники, тому їх було використано для подальшого аналізу.

3. Проведений аналіз білок-білкових взаємодій і встановлено взаємодію між білками людини серин-треонін протеїнкіназою MARK2 і 1- β -інтегрином для *cagA* і E-кадгерином для *fadA*.

4. Було проаналізовано сайти зв'язування цільових білків і визначено, що активними є 1 сайт у *cagA*, для *fadA* активні сайти зв'язування були відсутні, що відповідає як даним літератури, так і аналізу показників потенційних сайтів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–249.
2. Cancer [Електронний ресурс]: англійська версія / Режим доступу: <https://www.cancer.net/navigating-cancer-care/how-cancer-treated/personalized-and-targeted-therapies/what-targeted-therapy>
3. Doosey, C.M., Finn, K., Murphy, C. et al. The impact of the human microbiome in tumorigenesis, cancer progression, and biotherapeutic development. *BMC Microbiol* 22, 53 (2022).
4. Kim S, Covington A, Pamer EG. The intestinal microbiota: antibiotics, colonization resistance, and enteric pathogens. *Immunol Rev* 2017; P. 90-105.
5. Goodman, B., Gardner, H., 2018. The microbiome and cancer. *J. Pathol.*
6. Qin N, Zheng B, Yao J, et al. Influence of H7N9 virus infection and associated treatment on human gut microbiota. *Sci Rep* 2015;5: 14771.
7. Zhou, H., Beltrán, J.F. & Brito, I.L. Host-microbiome protein-protein interactions capture disease-relevant pathways. *Genome Biol* 23, 72 (2022).
8. Belcheva A, Irazabal T, Robertson SJ, et al. Gut microbial metabolism drives transformation of MSH2-deficient colon epithelial cells. *Cell* 2014.
9. Seidler KA, Seidler NW. Role of extracellular GAPDH in *Streptococcus pyogenes virulence*. *Mo Med.* 2013;110(3):236–40.
10. Kumar Singh P, Kashyap A, Silakari O. Exploration of the therapeutic aspects of Lck: a kinase target in inflammatory mediated pathological conditions. *Biomed Pharmacother.* 2018;108:1565–71.
11. Wolf AM, Wolf D, Rumpold H, Ludwiczek S, Enrich B, Gastl G, et al. The kinase inhibitor imatinib mesylate inhibits TNF- α production in vitro

and prevents TNF-dependent acute hepatic inflammation. *Proc Natl Acad Sci U S A*. 2005;102(38):13622–7.

12. Xu J, Liang R, Zhang W, Tian K, Li J, Chen X, et al. *Faecalibacterium prausnitzii*-derived microbial anti-inflammatory molecule regulates intestinal integrity in diabetes mellitus mice via modulating tight junction protein expression. *J Diabetes*. 2020;12(3):224–36.

13. Goodman B, Gardner H. The microbiome and cancer. *J Pathol*. 2018;244(5):667–76.

14. Chang WL, Yeh YC, Sheu BS. The impacts of *H. pylori* virulence factors on the development of gastroduodenal diseases. *J Biomed Sci*. 2018;25(1):1–9.

15. Kwok T, Zabler D, Urman S, Rohde M, Hartig R, Wessler S, et al. *Helicobacter* exploits integrin for type IV secretion and kinase activation. *Nature*. 2007;449:862–6.

16. Neal JT, Peterson TS, Kent ML, Guillemin KH. *pylori* virulence factor CagA increases intestinal cell proliferation by Wnt pathway activation in a transgenic zebrafish model. *DMM Dis Model Mech*. 2013;6(3):802–10.

17. Yamaoka Y, Kwon DH, Graham DY. A Mr 34,000 proinflammatory outer membrane protein (oipA) of *Helicobacter pylori*. *Proc Natl Acad Sci U S A*. 2000;97(13):7533–8.

18. Copenhagen-Glazer S, Sol A, Abed J, Naor R, Zhang X, Han YW, et al. Fap2 of *Fusobacterium nucleatum* is a galactose-inhibitable adhesin involved in coaggregation, cell adhesion, and preterm birth. *Infect Immun*. 2015;83(3):1104–13.

19. Abed J, Maalouf N, Parhi L, Chaushu S, Mandelboim O, Bachrach G. Tumor targeting by *Fusobacterium nucleatum*: A Pilot Study and future Perspectives. *Front Cell Infect Microbiol*. 2017;7(JUN):1–5.

20. Buc E, Dubois D, Sauvanet P, Raisch J, Delmas J, Darfeuille-Michaud A, et al. High Prevalence of Mucosa-Associated *E. coli* Producing Cyclomodulin and Genotoxin in Colon Cancer. *PLoS One*. 2013;8(2):e56964.

21. Dejea CM, Fathi P, Craig JM, Boleiji A, Taddese R, Geis AL, et al. Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science*. 2018;359(6375):592–7.
22. Dalmaso G, Cougnoux A, Delmas J, Darfeuille-Michaud A, Bonnet R. The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes*. 2015;5(5):675–80.
23. Sedighi M, Zahedi Bialvaei A, Hamblin MR, Ohadi E, Asadi A, Halajzadeh M, et al. Therapeutic bacteria to combat cancer; current advances, challenges, and opportunities. *Cancer Med*. 2019;8(6):3167–81.
24. Dash, S, Duraivelan, K, Samanta, D. Cadherin-mediated host–pathogen interactions. *Cellular Microbiology*. 2021; 23:e13316.
25. RCSB PDB [Електронний ресурс]: англійська версія / Режим доступу:
https://files.rcsb.org/pub/pdb/validation_reports/dv/4dvz/4dvz_full_validation.pdf
26. RCSB PDB [Електронний ресурс]: англійська версія / Режим доступу:
https://files.rcsb.org/pub/pdb/validation_reports/et/3etw/3etw_full_validation.pdf
27. Kumar, B. K., Faheem, Sekhar, K., Ojha, R., Prajapati, V. K., Pai, A., & Murugesan, S. (2022). Pharmacophore based virtual screening, molecular docking, molecular dynamics and MM-GBSA approach for identification of prospective SARS-CoV-2 inhibitor from natural product databases. *Journal of biomolecular structure & dynamics*, 40(3), 1363–1386.