

УНИВЕРЗИТЕТ У БЕОГРАДУ

Анђела М. Родић

БИОФИЗИЧКО МОДЕЛОВАЊЕ
РЕСТРИКЦИОНО-МОДИФИКАЦИОНИХ
СИСТЕМА БАКТЕРИЈА

докторска дисертација

Београд, 2022

UNIVERSITY OF BELGRADE

Andela M. Rodić

**BIOPHYSICAL MODELING OF BACTERIAL
RESTRICTION-MODIFICATION SYSTEMS**

Doctoral Dissertation

Belgrade, 2022

Ментори:

- Проф. др Марко Ђорђевић, ванредни професор, Универзитет у Београду – Биолошки факултет
- Др Магдалена Ђорђевић, научни саветник, Универзитет у Београду – Институт за физику

Чланови комисије:

- Проф. др Мирослав Живић, ванредни професор, Универзитет у Београду – Биолошки факултет
- Др Ксенија Радотић Хаџи-Манић, научни саветник, Универзитет у Београду, Институт за мултидисциплинарна истраживања
- Др Игор Салом, виши научни сарадник, Универзитет у Београду, Институт за физику

Датум одбране: 19.07.2022.

Изјаве захвалности

Желела бих да изразим највећу захвалност свом ментору, проф. др Марку Ђорђевићу, за несебичну подршку мом раду и развоју каријере у виду: дате прилике да учествујем у добро осмишљеним и изазовним истраживањима као члан неколико његових пројеката; подељених стручних савета, објашњења, одговора и вредних сазнања стечених богатим радним искуством и усавршавањем у Америци; стрпљиво вођених дискусија уз отвореност за моје идеје и недоумице; обезбеђених прилика за учешће на бројним научним скуповима, као и помоћи приликом пријављивања за учешће, стипендије и награде... Велико хвала за стварање и одржавање средине препуне изазова, али и награда за труд, повољне за развој мојих специфичних потенцијала, као и стручног знања и вештина, у којој сам стекла непроцењиво искуство, како за посао, тако и за живот.

Велику захвалност дугујем и менторки, др Магдалени Ђорђевић, за праћење и критичку анализу мог рада, као и бројне савете и предлоге који су допринели квалитету објављених радова. Високо ценим прилику да сарађујем са једном тако признатом и успешном научницом и да од ње учим.

Изузетан допринос истраживању приказаном у мојој докторској дисертацији је дала истраживачка група проф. др Константина Северинова која је извела лабораторијске експерименте који су мотивисали и допунили нашу теоријску анализу. Поносна сам на чињеницу да су резултати нашег заједничког труда два пута објављени у високоцитираном међународном часопису највише научне категорије, *Nucleic acids research*.

Захваљујем се проф. др Евгенију Здобнову на сарадњи у истраживању, као и др Бојани Илић, на сарадњи у послу, али и на дружењу у слободно време.

Ово истраживање је финансијски подржано пројектом Министарства просвете, науке и технологије Републике Србије под бројем ON173052, SCOPES пројектом IZ73Z0_152297 Швајцарске националне научне фондације и Међународним реинтеграционим грантом „Марија Кири“ под бројем PIRG08-GA-2010-276996.

Користим прилику да се осврнем на заслуге бројних људи из свог живота, који су ми својим делима помогли да се данас нађем пред одбраном доктората, и споменем макар неке од њих. Неизмерну захвалност дугујем својим родитељима, Мирославу и Надици Родић, који су ме научили разлици између добра и зла, достојанству и човечности као највишим моралним вредностима, као и тајни достизања моћног осећања испуњености услед свести о сопственој сврси и смислу, кроз постављање високих циљева и храбро улагање љубави, рада, времена и свих својих потенцијала у њихово постизање. Захваљујем се својој браћи, Растку и Михаилу, као и бабама – Младени и Радмили – и дедама – Гвоздену и Миливоју – за безусловну љубав и подршку. Хвала свим блиским, познатим и мање познатим рођацима који ме подразумевају као члана своје фамилије и то показују у критичним тренуцима. Велико хвала мојим пријатељима који су ми „доделили константну вредност“ у својим животима – Биљани Дудуковић, Бојани Рајковић, Кристини Сладоје и Јовани Деспотовић – као и свим другим драгим људима чија ми пажња и пријатељство обогаћују дане. На крају, желела бих да истакнем заслуге особа које су се игром случаја нашле у позицији мојих наставника, а које су из личних уверења преузеле на себе одговорност да ми помогну да развијем вредне потенцијале које су препознале у мени. Међу њима су свакако моја професорка клавира, Јелена Филотић, и професори Пете београдске гимназије, Милеса Анђус, Милица Шаренац и Ненад Јовановић. Ову Захвалницу завршавам споменувши само мали део људи који заслужују да се у њој нађу, са надом да сам успела да развејем све заблуде о томе да је докторска дисертација подухват за једног човека.

Биофизичко моделовање рестрикционо-модификационих система бактерија

Сажетак Рестрикционо-модификациони (P-M) и CRISPR-Cas системи користе различите механизме за обављање основне функције – одбране прокариотске ћелије од стране ДНК. За четири одабрана P-M система Типа II и CRISPR-Cas Типа I-E су постављени термодинамички модели регулације транскрипције и динамички модели експресије транскриптата и протеина. Симулацијом и анализом динамике модела су идентификоване особине динамике експресије система по покретању њихове активности у ћелији које вероватно представљају принципе еволутивног дизајна њихове регулације. Прецизније, испитано је: *i*) како пертурбације карактеристичних регулаторних својстава P-M система AhdI и EcoRV утичу на три предложена динамичка принципа; *ii*) да ли P-M систем Kpn2I, са регулацијом на нивоу елонгације транскрипције, може да обезбеди очекивана динамичка својства; *iii*) да ли су постојећа сазнања о регулацији P-M система Esp1396I довољна за репродуковање динамике експресије протеина измерене на нивоу појединачних ћелија; *iv*) какве особине вероватно има непозната динамика експресије CRISPR-Cas система у *Escherichia coli*, предвиђена уз претпоставку да се његов механизам регулације транскрипције може апроксимирати концептуално сличним механизмом P-M система. Показано је да сва четири P-M система, као и CRISPR-Cas, структурно испуњавају услове за постизање два предложена динамичка принципа – почетно кашњење експресије рестрикционе ендонуклеазе за експресијом метилтрансферазе и њен нагли пораст ка стационарном стању, док анализа система AhdI и EcoRV подржава и трећи – ниске флукуације у стационарном стању. Ова сазнања о дизајну природних генских мрежа омогућавају боље разумевање везе између њихове структуре и функције и дају смернице за дизајн синтетичких генских кола.

Кључне речи Рестрикционо-модификациони системи; CRISPR-Cas; Прокариотски имунски системи; Термодинамички модели; Динамички модели; Регулација транскрипције; Динамика експресије гена; Регулаторне генске мреже; Принципи дизајна

Научна област Биолошке науке

Ужа научна област Биофизика

Biophysical modeling of bacterial restriction-modification systems

Abstract Restriction-modification (R-M) and CRISPR-Cas systems use different mechanisms to perform their main function - defend prokaryotic cells from foreign DNA. Thermodynamic models of transcription regulation and dynamic models of transcript and protein expression were set for four selected Type II R-M systems and a Type I-E CRISPR-Cas. By simulating and analyzing the model dynamics, we identified the properties of the system expression dynamics upon the induction in a cell which may be the principles of the regulation evolutionary design. Specifically, we examined: i) how perturbing of the characteristic regulatory features of the R-M systems AhdI and EcoRV affects the three proposed dynamic principles; ii) if the R-M system Kpn2I, which regulation at the level of transcription elongation, can provide the expected dynamic properties; iii) if the known regulation of the R-M system Esp1396I is sufficient to reproduce the protein expression dynamics measured on single-cells; iv) which properties are probably found in the unknown expression dynamics of the CRISPR-Cas system in *Escherichia coli*, predicted under the assumption that its transcription regulation mechanism can be approximated by a similar one from R-M systems. We showed that all four R-M systems, as well as CRISPR-Cas, are able to achieve the two proposed dynamic principles - initial delay of restriction endonuclease with respect to methyltransferase expression and its rapid increase towards steady-state, while analysis of AhdI and EcoRV adds the third principle - low fluctuations in the steady-state. Gained insights into the design of these natural gene networks provide a better understanding of the relationship between their structure and function, as well as guidelines for the design of synthetic gene circuits.

Key words Restriction-modification systems; CRISPR-Cas; Prokaryotic immune systems; Thermodynamic models; Dynamic models; Transcription regulation; Gene expression dynamics; Regulatory gene networks; Design principles

Scientific field Life Sciences

Scientific subfield Biophysics

Садржај

1	УВОД	1
1.1	РЕСТРИКЦИОНО-МОДИФИКАЦИОНИ СИСТЕМИ	1
1.2	РЕГУЛАЦИЈА ЕКСПРЕСИЈЕ ОДАБРАНИХ Р-М СИСТЕМА ТИПА II	5
1.2.1	<i>Регулација експресије Р-М система AhdI</i>	6
1.2.2	<i>Регулација експресије Р-М система Esp1396I</i>	7
1.2.3	<i>Регулација експресије Р-М система EcoRV</i>	8
1.2.4	<i>Регулација експресије Р-М система Kpn2I</i>	8
1.3	СИСТЕМИ CRISPR-CAS	9
1.3.1	<i>Регулација експресије система CRISPR-Cas Tuna I-E</i>	9
1.4	ПРИНЦИПИ ТЕРМОДИНАМИЧКОГ МОДЕЛОВАЊА РЕГУЛАЦИЈЕ ТРАНСКРИПЦИЈЕ	11
1.4.1	<i>Извођење статистичких тежина полазећи од равнотежних биохемијских реакција</i>	11
1.4.2	<i>Извођење статистичких тежина према теорији статистичке термодинамике</i>	12
1.5	ПРИНЦИПИ ДИНАМИЧКОГ МОДЕЛОВАЊА ЕКСПРЕСИЈЕ ПРОТЕИНА	14
1.5.1	<i>Експериментално одређивање динамичке криве експресије протеина Р-М система у појединачној ћелији</i>	15
2	ХИПОТЕЗЕ И ЦИЉЕВИ ИСТРАЖИВАЊА	17
3	МЕТОДЕ	20
3.1	БИОФИЗИЧКО МОДЕЛОВАЊЕ Р-М СИСТЕМА AhdI	20
3.1.1	<i>Термодинамичко моделовање регулације транскрипције гена система AhdI</i>	20
3.1.2	<i>Динамичко моделовање експресије транскриптата и протеина система AhdI</i>	23
3.1.3	<i>Одређивање вредности параметара модела Р-М система AhdI</i>	24
3.2	БИОФИЗИЧКО МОДЕЛОВАЊЕ Р-М СИСТЕМА EcoRV	25
3.2.1	<i>Термодинамичко моделовање регулације транскрипције гена система EcoRV</i>	25
3.2.2	<i>Динамичко моделовање експресије транскриптата и протеина система EcoRV</i>	28
3.2.3	<i>Рескалирање величина и одређивање вредности параметара модела Р-М система EcoRV</i>	29
3.3	МЕРЕЊЕ ЕФЕКТА РЕМЕЋЕЊА СПЕЦИФИЧНИХ КОНСТРУКТИВНИХ КАРАКТЕРИСТИКА Р-М СИСТЕМА AhdI И EcoRV НА СВОЈСТВА ДИНАМИКЕ ЕКСПРЕСИЈЕ	30
3.3.1	<i>Дефинисање квантитативних мера за степен испољености жељених својстава динамике експресије Р-М система</i>	30
3.3.2	<i>Пертурбисање равнотежне константе дисоцијације за реакцију димеризације С протеина</i>	33
3.3.3	<i>Пертурбисање нивоа кооперативности у везивању два С димера за ДНК и константе транслације за С протеин</i>	34
3.4	БИОФИЗИЧКО МОДЕЛОВАЊЕ Р-М СИСТЕМА Kpn2I	34
3.4.1	<i>Термодинамичко моделовање регулације транскрипције гена система Kpn2I</i>	34
3.4.2	<i>Динамичко моделовање експресије транскриптата и протеина система Kpn2I</i>	37
3.4.3	<i>Рескалирање величина и одређивање вредности параметара модела Р-М система Kpn2I</i>	38
3.5	БИОФИЗИЧКО МОДЕЛОВАЊЕ Р-М СИСТЕМА Esp1396I	39
3.5.1	<i>Термодинамичко моделовање регулације транскрипције гена система Esp1396I</i>	39
3.5.2	<i>Динамичко моделовање експресије транскриптата и протеина система Esp1396I</i>	42
3.5.3	<i>Одређивање вредности параметара модела Р-М система Esp1396I фитовањем података из експерименталног мерења динамике експресије протеина у ћелији</i>	42
3.6	БИОФИЗИЧКО МОДЕЛОВАЊЕ СИСТЕМА CRISPR-CAS	44

3.6.1	Динамичко моделовање експресије RNK и протеина синтетичких генских кола базираних на систему CRISPR-Cas	45
3.6.2	Термодинамичко моделовање регулације транскрипције у „конститутивном моделу“ експресије система CRISPR-Cas	46
3.6.3	Термодинамичко моделовање регулације транскрипције у „кооперативном моделу“ експресије система CRISPR-Cas	48
3.6.4	Одређивање и подешавање вредности параметара модела система CRISPR-Cas	49
4	РЕЗУЛТАТИ	52
4.1	ЕФЕКТИ РЕМЕЋЕЊА СПЕЦИФИЧНИХ КОНСТРУКТИВНИХ КАРАКТЕРИСТИКА Р-М СИСТЕМА АНДI НА СВОЈСТВА ДИНАМИКЕ ЊЕГОВЕ ЕКСПРЕСИЈЕ	52
4.1.1	Повећавање афинитета С протеина за формирање хомодимера	52
4.1.2	Повећавање ефикасности транслације транскрипта <i>ahdIC</i>	53
4.1.3	Смањивање афинитета за кооперативно везивање два димера С протеина	54
4.2	ПРЕДВИЂАЊЕ ДИНАМИКЕ ЕКСПРЕСИЈЕ Р-М СИСТЕМА ЕСОРV	55
4.3	ЕФЕКТИ УВОЂЕЊА СПЕЦИФИЧНИХ КОНСТРУКТИВНИХ КАРАКТЕРИСТИКА Р-М СИСТЕМА АНДI У Р-М СИСТЕМ ЕСОРV НА СВОЈСТВА ДИНАМИКЕ ЊЕГОВЕ ЕКСПРЕСИЈЕ	56
4.3.1	Смањивање афинитета С протеина за формирање хомодимера	56
4.3.2	Смањивање ефикасности транслације с транскрипта	57
4.3.3	Повећавање афинитета за кооперативно везивање два димера С протеина	58
4.4	ПРЕДВИЂАЊЕ ДИНАМИКЕ ЕКСПРЕСИЈЕ Р-М СИСТЕМА КРН2I И АНАЛИЗА УЛОГЕ РЕГУЛАТОРНОГ МЕХАНИЗМА БЛОКИРАЊА ЕЛОНГАЦИЈЕ ТРАНСКРИПЦИЈЕ	59
4.5	ПРЕДВИЂАЊЕ ДИНАМИКЕ ЕКСПРЕСИЈЕ Р-М СИСТЕМА ЕСП1396I И АНАЛИЗА НАЈБОЉЕГ ФИТА ЕКСПЕРИМЕНТАЛНИХ ПОДАТАКА	61
4.6	ЕФЕКТИ СПЕЦИФИЧНИХ КОНСТРУКТИВНИХ КАРАКТЕРИСТИКА СИСТЕМА CRISPR-CAS НА СВОЈСТВА ДИНАМИКЕ ЊЕГОВЕ ЕКСПРЕСИЈЕ	62
4.6.1	Ефекти постепене синтезе протеина Casбе на динамику експресије <i>pre-crRNK</i> и <i>crRNK</i>	63
4.6.2	Ефекти смањивања стопе неспецифичне деградације <i>pre-crRNK</i> на динамику експресије <i>pre-crRNK</i> и <i>crRNK</i>	66
4.6.3	Ефекти повећавања транскрипционе активности промотора P.CRISPR на динамику експресије <i>pre-crRNK</i> и <i>crRNK</i>	67
5	ДИСКУСИЈА	70
6	ЗАКЉУЧАК	75
7	ЛИТЕРАТУРА	78

1 Увод

Прокариотски организми поседују две врсте распрострањених молекуларних система за одбрану ћелије од унетих страних молекула ДНК, који обезбеђују њихово препознавање и разликовање од ДНК генома домаћина, као и њихово елиминисање (Goldberg & Marraffini, 2015). Рестрикционо-модификациони системи представљају једноставнији тип прокариотских одбрамбених система, који се заснива на стратегији модификације специфичних секвенци у геному ћелије и уништавања само ДНК молекула који садрже такву немодификовану секвенцу. Овакав механизам не обезбеђује високу специфичност одбране и могућност памћења ранијих инфекција, па прокариоти имају и напредну врсту одбрамбених система, CRISPR-Cas, која им даје ове способности. Главни објект истраживања представљеног у овој докторској дисертацији су рестрикционо-модификациони системи Типа II у бактеријама, који свој релативно једноставан задатак обављају користећи зачуђујуће велики број разноврсних механизма регулације својих гена. Проучавање логике еволутивног дизајна њихових регулаторних механизма може да омогући боље разумевање регулације експресије природних генских мрежа, тј. начина на који се комбинаторно граде генетички системи жељеног понашања. У овој докторској дисертацији су представљене процедуре биофизичког моделовања регулације четири одабрана рестрикционо-модификациона система, као и резултати анализе симулираног понашања модела. Предложена су конкретна жељена својства понашања ове класе система која могу представљати принципе дизајна њихове регулације, тј. њихову заједничку основу упркос запаженој конструктивној разноврсности. Као корак даље у тражењу универзалних принципа, задатих потребама функције система, у основи различитих стратегија, постулирана је сличност општих особина механизма регулације експресије у рестрикционо-модификационом и CRISPR-Cas систему. Та сличност је искоришћена за предвиђање особина понашања CRISPR-Cas система чију динамику експресије *in vivo* још увек није могуће експериментално окарактерисати.

1.1 Рестрикционо-модификациони системи

Рестрикционо-модификациони (Р-М) системи свој назив дугују пару међусобно супротстављених функција које обављају у прокариотској ћелији (Ershova et al., 2015; Rodic et al., 2018). Њихове основне компоненте – рестрикциона ендонуклеаза и метилтрансфераза – специфично препознају исте ДНК секвенце. Ако ДНК секвенцу прва препозна рестрикциона ендонуклеаза, она ће својом ендонуклеазном активношћу увести прекид у дволанчани молекул ДНК унутар те секвенце. Ако пак метилтрансфераза модификује дату ДНК секвенцу додавањем метил група, секвенца ће бити заштићена од дејства рестрикционе ендонуклеазе.

Р-М системи су широко распрострањени код бактерија, а присутни су и код археа (Ershova et al., 2015; Makarova et al., 2013; M. O. Nagornykh et al., 2008; A. Pingoud et al., 2014; Roberts et al., 2005, 2015; Vasu & Nagaraja, 2013). Подељени су на Типове према структури компоненти које су носиоци основних ензимских активности система (Ershova et al., 2015; Mruk & Kobayashi, 2014). У системима Типа I засебно кодиране протеинске субјединице за специфично препознавање и модификацију секвенце формирају комплекс који метилује ДНК, док активност рестрикционе ендонуклеазе обавља комплекс који поред ове две садржи и субјединицу за рестрикцију. Слично, у системима Типа III пресецање ДНК такође обавља комплекс који се састоји из протеина за рестрикцију и протеина за модификацију, док сам протеин за модификацију има активност метилтрансферазе. Предмет изучавања ове докторске

дисертације представљају одабрани Р-М системи Типа II у којима одвојени протеини носе ензимске активности рестрикционе ендонуклеазе и метилтрансферазе, којима делују на циљне, палиндромске секвенце ДНК дуге 4-8 базних парова (Ershova et al., 2015; M. O. Nagornykh et al., 2008; A. Pingoud et al., 2014; Roberts et al., 2005; Wilson & Murray, 1991). Постоји и Тип IV, у који се убрајају Р-М системи који се састоје од рестрикционе ендонуклеазе која специфично пресеца метиловане циљне секвенце и вероватно има улогу у изазивању смрти ћелије са конфликтним обрасцима метилације које намећу различити, истовремено присутни Р-М системи (Mruk & Kobayashi, 2014).

Функционалне компоненте Р-М система Типа II су кодиране малим, компактним кластерима гена који се, за разлику од Типова I и III, у природи најчешће налазе на потенцијално мобилним генетичким елементима, као што су плазмиди широког спектра домаћина (Furuta & Kobayashi, 2013; Jeltsch & Pingoud, 1996; Karyagina et al., 1997; Kobayashi, 2001, 2004; Makarova et al., 2011; Mruk & Kobayashi, 2014; C. Price & Bickle, 1986). Путем хоризонталног трансфера гена у бактеријској популацији, ћелије могу задобити функционалности нових Р-М система ако приме одговарајуће плазмиде. У складу са оваквим начином разношења, највише Р-М система је пронађено у бактеријама које су природно склоне усвајању ДНК молекула из спољашње средине, као што су *Helicobacter pylori*, *Neisseria gonorrhoeae*, *Neisseria meningitidis* и *Haemophilus influenzae*, док паразитским бактеријама које живе унутар ћелија домаћина, као што су *Mycoplasma*, *Buchnera*, *Treponema* и *Chlamydia*, они вероватно не доносе селективну предност, будући да их оне углавном немају (Mruk & Kobayashi, 2014).

Бројне улоге Р-М система (Vasu & Nagaraja, 2013) и начини на које усмеравају еволуцију бактеријских популација (Furuta & Kobayashi, 2012) сведоче о њиховом великом биолошком значају (Ershova et al., 2015; Vasu & Nagaraja, 2013). Својим деловањем они стварају баријеру за проток гена између ћелијских линија са различитим епигенетичким, тј. обрасцима метилације генома, чиме потпомажу акумулацију разлика и раздвајање сојева (Ershova et al., 2015; Mruk & Kobayashi, 2014). Обрасци метилације, које може успоставити и више метилтрансфераза које припадају различитим Р-М системима, мењају глобалну експресију гена ћелије, при чему су изменама нарочито подложни гени у непосредној близини метилованих циљних секвенци. Разлике у експресији гена могу резултирати и различитим ћелијским фенотиповима који патогеним бактеријама могу донети пожељне особине вируленције, покретљивости, производње протеина топлотног удара и др. (Casadesús & Low, 2006; Ershova et al., 2015; Mruk & Kobayashi, 2014; Srikhanta et al., 2011). Додатно, резистенција на бактериофаге за коју су заслужни Р-М системи бактеријама даје предност у колонизацији нових станишта (Ershova et al., 2015; Korona & Levin, 1993). Одбрана бактерије од унете стране ДНК се сматра основном функцијом Р-М система (Vasu & Nagaraja, 2013). Они се називају и рудиментарним прокариотским „имунским“ системима, по аналогији са имунским системима вишећелијских еукариота, јер поседују механизме за разликовање сопствене од стране ДНК, заштиту сопствене, и специфично уништавање само стране ДНК (Ershova et al., 2015). Наиме, главни принцип деловања Р-М система подразумева да се циљне секвенце у геному домаћина модификују метилацијом, а да рестрикциона ендонуклеаза препознаје и сече само циљне секвенце унете ДНК које су немодификоване (Bogdanova et al., 2008; Gingeras & Brooks, 1983; Kiss et al., 1985; Kobayashi, 2001; Semenova et al., 2005).

У прилог имунске функције Р-М система говоре бројна запажања из биологије и еволуције бактериофага, на које бактерије могу постати резистентне захваљујући овим системима. За неколико Р-М система је експериментално показано да се у њиховом присуству смањује ефикасност формирања вирусних плака за неколико редова величине (Bogdanova et al., 2009; Gingeras & Brooks, 1983; Kiss et al., 1985; M. Nagornykh et al., 2011; Semenova et al., 2005). Да је дејство Р-М система код бактерија извршило еволуциони притисак на бактериофаге (Krüger & Bickle, 1983; Tock & Dryden, 2005) показују механизми које су различити бактериофаги развили како би их избегли (Korona & Levin, 1993; Makino et al., 1980; Morozova et al., 2016;

Takahashi & Marmur, 1963; Tock & Dryden, 2005). На пример, неки бактериофаги у свом геному садрже гене за анти-рестрикциони протеин (Makino et al., 1980) или за ензиме који заштитно модификују секвенце у њиховом геному (Takahashi & Marmur, 1963). Неки бактериофаги се служе алтернативним базама, као што је уридин, у својој ДНК како би избегли препознавање од стране рестрикционе ендонуклеазе (Takahashi & Marmur, 1963). Многи бактериофаги у свом геному не садрже циљне секвенце које специфично препознају Р-М системи распрострањени у бактеријама врсте или соја који инфицирају, јер су их временом изгубили (Bickle & Krüger, 1993; Kawamura et al., 1981; Korona et al., 1993; Rocha et al., 2001).

Будући да кодирају ензиме који у ћелији имају улоге отрова (рестрикциона ендонуклеаза која може да уништи геном домаћина) и антиотрова (метилтрансфераза која штити геном), балансирана експресија ових ензима је услов за преживљавање ћелије домаћина (Mruk & Kobayashi, 2014). Предложено је да се, због тога, Р-М системи понашају као „себични“ генетички елементи или модули који „изазивају зависност“ (Kobayashi, 2001). Конкретно, ако ћелија приликом деобе изгуби плазмид са генима Р-М система, може доћи до њене програмиране смрти (енг. *postsegregational killing*) услед нарушавања интегритета ДНК дејством токсичне рестрикционе ендонуклеазе, која понекад има и дужи полу-живот од метилтрансферазе (Asakura & Kobayashi, 2009; Handa & Kobayashi, 1999; Ichige & Kobayashi, 2005; Kulakauskas et al., 1995; Mruk & Kobayashi, 2014; Naito et al., 1995; Ohno et al., 2008). Промене транскриптома које се притом јављају су сличне онима које се уочавају у ћелијама третираним бактериоцидним антибиотицима (Asakura & Kobayashi, 2009). Постсегрегацијско убијање је карактеристично само за Р-М системе Типа II који су, за разлику од осталих Типова, изразито мобилни. Овај ефекат, који ћелију чини зависном од гена Р-М система, фаворизује њихово очување кроз генерације (Mruk & Kobayashi, 2014).

Из свега наведеног јасно следи да експресија протеина Р-М система у ћелији мора бити строго контролисана, при чему је за ћелију нарочито изазован тренутак првог сусрета са новим Р-М системом, тј. уласка плаزمида који кодира његове елементе (М. О. Nagornykh et al., 2008). Бројна експериментална запажања говоре у прилог претпоставци да је за безбедно успостављање Р-М система у новој ћелији домаћина (а и за опстанак самог плазмида) кључно кашњење појаве рестрикционе ендонуклеазе у ћелији за акумулацијом метилтрансферазе (Handa et al., 2000; Ichige & Kobayashi, 2005; Katna et al., 2010; Morozova et al., 2016; Mruk & Blumenthal, 2008; М. О. Nagornykh et al., 2008). Међутим, кашњење је директно демонстрирано регистравањем динамике експресије протеина само једног Р-М система Типа II, PvuII (Mruk & Blumenthal, 2008), а подржано је и праћењем транскрипционе активности промотора и активности ензима Р-М система Типа I, hsdK (Prakash-Cheng et al., 1993; Prakash-Cheng & Ryu, 1993). Ово кашњење треба да обезбеди довољно времена да метилтрансфераза модификује циљне секвенце у геному домаћина и на тај начин их заштити од сечења. Експериментално је показано да се при недовољној експресији метилтрансферазе, у геному бактерије повећава број немодификованих циљних секвенци које временом постају мета рестрикционе ендонуклеазе и ћелија умире (Ichige & Kobayashi, 2005). Према томе, по уласку плазмида у ћелију метилтрансфераза треба да се синтетише у релативно великој количини. Међутим, очекује се да касније током времена њена количина у ћелији опадне, како би се смањила могућност заштитне модификације стране ДНК. Показани примери који овакво очекивање подржавају се односе на бактериофаге чија је ДНК модификована од стране метилтрансферазе бактеријске ћелије, и чије даље генерације једнако ефикасно инфицирају ћелије са и без Р-М система (Bogdanova et al., 2009; Enikeeva et al., 2010; Gingeras & Brooks, 1983; Kiss et al., 1985; Klimuk et al., 2018; Morozova et al., 2016; М. Nagornykh et al., 2011; Semenova et al., 2005).

Поред самих гена Р-М система, на плазмиду се налази и комплетна генетичка информација неопходна за регулацију њихове експресије у ћелији домаћина, што систем чини независним од транскрипционих фактора домаћина и њиховог тренутног нивоа у ћелији (М. О. Nagornykh et al., 2008). С обзиром на релативно једноставан механизам деловања Р-М система, делује

изненађујуће откриће да се они у природи могу наћи са разноврсним регулаторним елементима, компонентама и својствима која комбинаторно граде различите механизме регулације експресије гена (М. О. Nagornykh et al., 2008; А. Pingoud et al., 2014; Rodic, Vladojević, Zdobnov, et al., 2017). На пример, на нивоу организације гена за рестрикциону ендонуклеазу и метилтрансферазу на плазмиду, могу се разликовати Р-М системи са конвергентном, дивергентном и линеарном архитектуром (Bogdanova et al., 2008; Karyagina et al., 1997; Morozova et al., 2016; М. О. Nagornykh et al., 2008; Semenova et al., 2005; Sohail et al., 1995; Sorokin et al., 2009; Tao et al., 1991; Zakharova et al., 2004). У конвергентној архитектури, промотори ових гена су усмерени један према другом, а у дивергентној супротно и делимично су преклопљени. Чести су и системи са линеарном архитектуром, у којима су гени за рестрикциону ендонуклеазу и метилтрансферазу сложени један за другим, а могу чинити и оперон (Mruk & Kobayashi, 2014; М. О. Nagornykh et al., 2008).

Велика група Р-М система садржи додатни ген за тзв. контролни (С) протеин који има улогу транскрипционог фактора у регулацији експресије једног или оба гена који кодирају ензиме (Anton et al., 1997; Ball et al., 2012; Bogdanova et al., 2008; Cesnaviciene et al., 2003; Ives et al., 1992; Knowle et al., 2005; Lubys et al., 1999; McGeehan et al., 2006; М. О. Nagornykh et al., 2008; Rezulak et al., 2016; Semenova et al., 2005; Sohail et al., 1995; Sorokin et al., 2009; Tao et al., 1991, 1991). С протеини су еволутивно сродни репресорима λ -бактериофага који су компоненте генетичког прекидача који ове вирусе преводи у лизогено стање (Ptashne, 2005; Sawaya et al., 2005; Stayrook et al., 2008). Кристалне структуре С протеина из појединих Р-М система Типа II показују да они формирају хомодимере, док су С мономерни структурно веома слични ДНК-везујућем домену репресора λ -бактериофага (McGeehan et al., 2005; Sawaya et al., 2005). ДНК секвенца за коју се С протеин везује у форми димера је аналогна оператору λ -бактериофага (М. О. Nagornykh et al., 2008). Два места за кооперативно везивање два С димера чине тзв. С-box (Bart et al., 1999; Bogdanova et al., 2008, 2009; Mruk et al., 2007; Sorokin et al., 2010). Ген за С протеин типично чини оперон са геном за рестрикциону ендонуклеазу (Anton et al., 1997; Bogdanova et al., 2008; Cesnaviciene et al., 2003; Kita et al., 2002; Rimseliene et al., 1995; Semenova et al., 2005; Vijesurier et al., 2000), а С-box се обично налази узводно од промотора оперона са којим је делимично преклопљен, мада се могу наћи и разне друге конфигурације секвенци. С протеини доводе до савијања ДНК молекула у региону за који су везани и, у зависности од конфигурације њихових места за везивање, могу утишавати транскрипцију гена, или је активирати интеракцијом са σ субјединицом РНК полимеразе (Mruk & Kobayashi, 2014; М. О. Nagornykh et al., 2008; Sorokin et al., 2010).

Међу елементима регулације Р-М система се уочавају неки који могу допринети кашњењу експресије рестрикционе ендонуклеазе у ћелији и почетној акумулацији метилтрансферазе (М. О. Nagornykh et al., 2008). Прво, кашњење се генерише на нивоу базалне транскрипције гена, када у ћелији још нема ниједног продукта гена Р-М система, будући да је он тек ушао у њу на плазмиду. Промотор гена за метилтрансферазу је обично јачи од промотора гена за рестрикциону ендонуклеазу што омогућава интензивну почетну транскрипцију гена за метилтрансферазу (Bogdanova et al., 2008, 2009; Vijesurier et al., 2000). У случају дивергентних, преклопљених промотора, РНК полимеразе која се са већим афинитетом везује за промотор гена за метилтрансферазу стерно омета везивање РНК полимеразе за промотор гена за рестрикциону ендонуклеазу (Karyagina et al., 1997; Zakharova et al., 2004). Транскрипција са слабог промотора за рестрикциону ендонуклеазу се активира тек након достизања празног нивоа метилтрансферазе (Beletskaya et al., 2000; Christensen & Josephsen, 2004; Protsenko et al., 2009; Som & Friedman, 1997; Zakharova et al., 2004) или С протеина који има улогу активатора када се њихови гени налазе у оперону и преписују са заједничког промотора (Bogdanova et al., 2008, 2009).

Друго, синтеза рестрикционе ендонуклеазе може бити посредно додатно одложена успоравањем синтезе активаторског С протеина на нивоу транслације, одсуством јаког места

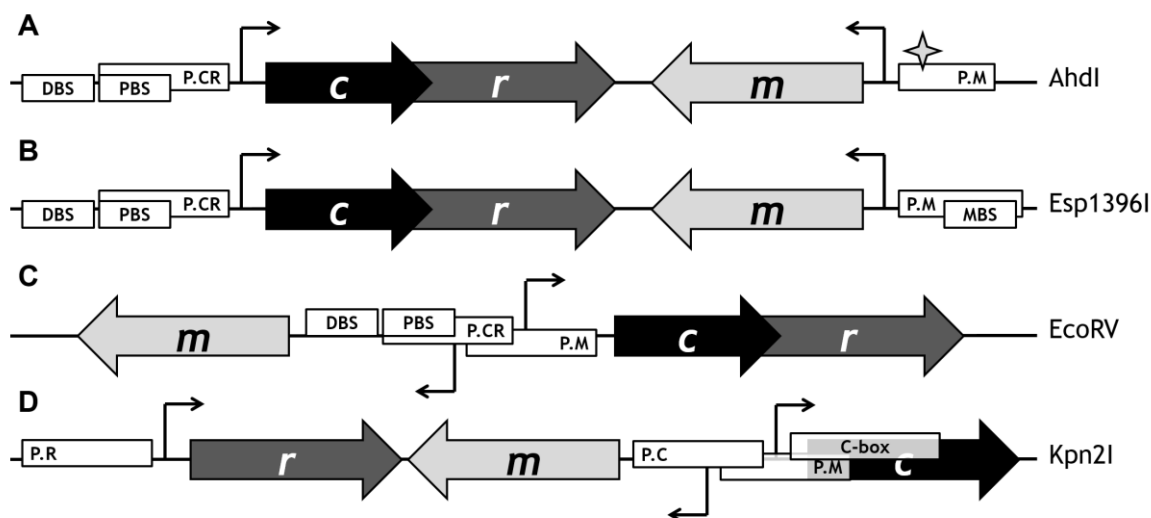
за везивање рибозома (Moll et al., 2002). Њена синтеза се може и директно одложити на нивоу транслације сакривањем места за везивање рибозома РНК укосницом и спрезањем транслације њеног цистрона са транслацијом узводног цистрона за С протеин (Cesnaviciene et al., 2003). На тај начин се обезбеђује да транслација низводног цистрона за рестрикциону ендонуклеазу може да почне тек кад се заврши транслација узводног цистрона за С протеин. У линеарним Р-М системима су често присутни промотори са којих се преписује антисенс ланац ДНК у региону оперона кодираним сенс ланцем. Антисенс РНК се вероватно комплементарно спарује са бицистронском информационом РНК, насталом транскрипцијом сенс ланца, чиме убрзава њену деградацију и онемогућава транслацију (Mruk et al., 2011; Mruk & Kobayashi, 2014; M. Nagornykh et al., 2011).

Треће, за разлику од метилтрансферазе, рестрикционе ендонуклеазе Типа II су обично ензимски активне у форми хомодимера, па се појава активне форме у ћелији одлаже за време потребно за достизање одређене пражне концентрације протеина (Ershova et al., 2015; A. Pingoud et al., 2005; Wilson, 1991). У неким системима, С протеин одликује релативно низак афинитет за димеризацију, што такође доприноси кашњењу будући да он активира транскрипцију у форми димера (Bogdanova et al., 2008).

Одлагањем синтезе рестрикционе ендонуклеазе се обезбеђује време током којег ниво метилтрансферазе у ћелији може да расте. Међутим, нарочито у каснијем периоду, тај ниво не би смео да постане превисок јер се тиме повећава могућност модификације стране ДНК. У неким случајевима, као што је у дивергентној архитектури система, експресија метилтрансферазе и директно спречава експресију рестрикционе ендонуклеазе (Karyagina et al., 1997; Zakharova et al., 2004). Због тога се у Р-М системима уочавају и регулаторна својства којима се смањује експресија метилтрансферазе. На пример, унутар промотора њеног гена се може налазити циљна секвенца коју метилтрансфераза препознаје и модификује. У метилованој форми, промотор постаје утишан. На овај начин систем има способност да осети и реагује на стање метилације у ћелији (Beletskaya et al., 2000; Bogdanova et al., 2008; Christensen & Josephsen, 2004; Kita et al., 1989; Lubys & Janulaitis, 1995; Zakharova et al., 2004). Метилтрансферазе неких Р-М система имају додатни ДНК-везујући домен и могу да делују као репресори транскрипције сопственог гена везујући се за оператор у региону промотора, независно од своје ензимске функције (Butler & Fitzgerald, 2001; Karyagina et al., 1997; Som & Friedman, 1994, 1997). С протеин такође може имати улогу репресора транскрипције гена за метилтрансферазу (Bogdanova et al., 2009). Метилтрансферазе типично имају краћи полу-живот од рестрикционих ендонуклеаза, па и та особина доприноси смањењу њихове концентрације у ћелији у каснијем периоду, када им је експресија утишана.

1.2 Регулација експресије одабраних Р-М система Типа II

У овој докторској дисертацији је моделована и анализирана регулација експресије четири Р-М система (Слика 1) које одликују различити регулаторни механизми и употребљени елементи, односно стратегије безбедног и ефикасног успостављања система у бактеријској ћелији.



Слика 1. Организација гена P-M система (A) AhdI, (B) Esp1369I, (C) EcoRV и (D) Kpn2I на плазмидима. Гени који кодирају C протеин (*c*), рестрикциону ендонуклеазу (*r*) и метилтрансферазу (*m*) су приказани широком стрелицама, док правоугаоници означени са P означавају позиције одговарајућих промотора. Танке стрелице указују на смер транскрипције са датих промотора. Правоугаоници означени са DBS, PBS, MBS и C-box означавају места за везивање C протеина, као што је објашњено у главном тексту. Звездица симболизује присуство циљне секвенце у региону P.M коју метилтрансфераза модификује. Слика модификована из (Rodic et al., 2018).

1.2.1 Регулација експресије P-M система AhdI

На основу распореда и оријентације експериментално мапираних гена и њихових промотора, AhdI се сврстава у P-M системе са конвергентном архитектуром (Слика 1A) (Bogdanova et al., 2008). Оперон који садржи гене за C протеин и рестрикциону ендонуклеазу се преписује са промотора који је овде означен са P.CR, док се ген за метилтрансферазу преписује са промотора P.M. Промотори су раздвојени генима система и усмерени су један према другом.

Показано је да се C протеин за ДНК везује у форми димера, а не мономера (Streeter et al., 2004). У региону P.CR, идентификоване су ДНК секвенце два везивна места за димере C протеина, тзв. C-box (Bogdanova et al., 2008; McGeehan et al., 2006), која су овде названа дистално (DBS, од енгл. *distal binding site*) и проксимално везивно место (PBS, од енгл. *proximal binding site*), при чему се PBS преклапа са -35 елементом промотора P.CR. Када је фрагмент ДНК који садржи C-box помешан са C протеинима присутним у различитим концентрацијама, раздвајањем формираних комплекса електрофорезом на гелу су у свим случајевима добијене (обе или једна) следеће траке: трака која одговара слободној ДНК, и трака која одговара мобилности тетрамера C протеина везаног за ДНК (Bogdanova et al., 2008). Будући да ни при изузетно ниским концентрацијама C протеина није уочена трака која одговара комплексу у коме је само један C димер на ДНК, закључено је да се два C димера везују за C box са веома високом кооперативношћу (McGeehan et al., 2006; Streeter et al., 2004). Како DBS има много већи афинитет за везивање C протеина од PBS (Bogdanova et al., 2008; McGeehan et al., 2006), које је превише слабо да би се C димер за њега независно везао, C димер се самостално везује за DBS, да би затим кооперативном интеракцијом омогућио веома брзо везивање другог C димера за PBS. Притом, показано је да C тетрамер везан за ДНК утишава транскрипцију сопственог и гена за рестрикциону ендонуклеазу. Ово запажање је указало на потребу за активацијом транскрипције са промотора P.CR, јер би у супротном по уласку система на плазмиду у ћелију ниском базалном активношћу P.CR била брзо достигнута пражна концентрација C протеина при којој би се репресорски комплекс формирао и спречио синтезу довољне количине рестрикционе ендонуклеазе (Bogdanova et al., 2008; McGeehan et al., 2006). Претпостављено је да C димер везан за DBS, осим другог C димера, на ДНК може да регрутује и РНК полимеразу и на тај начин активира транскрипцију. У прилог овој претпоставци,

идентификовани су аминокиселински остаци С протеина који могу учествовати у овим интеракцијама (Bogdanova et al., 2008).

Наша истраживачка група је раније поставила термодинамички модел регулације транскрипције са промотора Р.СР у Р-М систему AhdI (Bogdanova et al., 2008). Равнотежна транскрипциона активност промотора је експериментално измерена *in vitro* при различитим концентрацијама С протеина у: 1) изворном, *wild type* (*wt*) систему, као и у системима са уведеним мутацијама које смањују афинитет 2) DBS, 3) PBS, или 4) оба места за везивање С протеина (Bogdanova et al., 2008). У сва четири случаја је добијен добар фит криве предвиђене моделом на експерименталне податке, при релативним вредностима параметара које су биле очекиване на основу претпоставки о регулацији система. Како је за реакцију димеризације С протеина у систему AhdI одређена висока вредност равнотежне константе дисоцијације (2.5 mM) (Streeter et al., 2004), при којој су С протеини у раствору претежно у форми мономера, конструисан је и алтернативни модел који подразумева њихову претежно димерну форму. Значајно умањена моћ овог алтернативног модела да репродукује измерену зависност активности Р.СР од нивоа С протеина је указала на незанемарљив утицај ниског афинитета С протеина за димеризацију на одговор система AhdI. Тиме је такође подржан избор претпоставки о регулацији система које су уграђене у главни модел. Будући да он веродостојно описује регулацију транскрипције са Р.СР у систему AhdI, у овој докторској дисертацији је адаптиран за извођење анализа од интереса (Rodic, Blagojevic, Zdobnov, et al., 2017).

Још једно регулаторно својство карактеристично за Р-М систем AhdI је одсуство „водичке“ секвенце, 5'-UTR (енг. *UnTranslated Region*), са транскрипта оперона (енг. *leaderless transcript*), тако да полицистронска РНК директно почиње првим кодоном цистрона који кодира С протеин и тај цистрон нема узводно место за везивање рибозома. Овакви транскрипти су у *E. coli* много ређи од канонских и, мада неки подаци указују на то да се S70 субјединица рибозома може везати за њих независно и од иницијаторске транспортне РНК, њихова транслација је мање ефикасна (Laursen et al., 2005; O'Donnell & Janssen, 2001). Према томе, очекује се да је транслација цистрона за С протеин мање ефикасна од транслације цистрона за рестрикциону ендонуклеазу и метилтрансферазу.

Циљна секвенца коју препознају ензими система AhdI је идентификована у пределу -10 елемента промотора Р.М (Слика 1А). Метилтрансфераза може да регулише експресију свог гена негативном повратном спрегом, тј. може да метилује дату секвенцу чиме се утишава промотор Р.М (Bogdanova et al., 2008). Ова и остале доступне експерименталне информације о систему су искоришћене за конструисање једноставног динамичког модела експресије транскрипата и протеина система, који је такође адаптиран у овој докторској дисертацији. Симулацијом модела је предвиђена способност система AhdI да оствари кашњење експресије рестрикционе ендонуклеазе за експресијом метилтрансферазе у бактеријској ћелији (Bogdanova et al., 2008).

1.2.2 Регулација експресије Р-М система Esp1396I

Р-М систем Esp1396I је лоциран на плазмидима соја RFL1396 *Enterobacter* sp. (Cesnaviciene et al., 2003). Као и систем AhdI, одликује се конвергентном архитектуром (Слика 1В), као и истим обрасцем везивања С протеина за С-бох којим се регулише транскрипција оперона са генима за С протеин и рестрикциону ендонуклеазу. У овом систему је измерена нешто нижа јачина кооперативне интеракције два С димера на ДНК него у систему AhdI (Bogdanova et al., 2009). Посебна карактеристика овог система је присуство трећег места за везивање С протеина у региону промотора Р.М. Везивањем С димера за ово место се спречава везивање РНК полимеразе за промотор, тј. негативно се регулише транскрипција гена за метилтрансферазу (Bogdanova et al., 2009; Cesnaviciene et al., 2003). Познато је и да промотор Р.М у овом систему има виши афинитет за везивање РНК полимеразе него Р.СР. У погледу везивања С димера за ДНК, МBS је место са највећим афинитетом, затим следи DBS, док је PBS најслабије место

(Bogdanova et al., 2009). Још једно регулаторно својство описано на систему Esp1396I је спрегнутост транслације (енг. *coupled translation*) цистрона за С протеин и цистрона за рестрикциону ендонуклеазу на полицистронској информационој РНК. Наиме, место за везивање рибозома низводног цистрона за рестрикциону ендонуклеазу је сакривено унутар укоснице коју формира РНК. Ово место постаје доступно за везивање рибозома када се укосница расклопи, а за то је потребна активна (и комплетна) транслација узводног цистрона за С протеин.

1.2.3 Регулација експресије Р-М система EcoRV

Р-М систем EcoRV има дивергентну архитектуру (Слика 1С). Његови промотори, који се налазе у интергенском региону, као и места за везивање С протеина су експериментално мапирани (Semenova et al., 2005). Промотор Р.М, са којег се врши транскрипција гена за метилтрансферазу, и промотор Р.СR, са којег се преписује оперон у чијем су саставу гени за С протеин и рестрикциону ендонуклеазу, су међусобно преклопљени и окренути у супротним смеровима. За разлику од транскрипције оперона и транскрипције гена за метилтрансферазу у конвергентним системима, које су релативно независне једна од друге, овде су оне снажно међусобно спрегнуте (Rodic, Blagojevic, Zdobnov, et al., 2017). Везивање РНК полимеразе за Р.М и за Р.СR се међусобно искључују, при чему се у режиму базалне транскрипције РНК полимеразе преферентно везује за јачи, Р.М промотор истовремено утишавајући Р.СR (Semenova et al., 2005). Комплетан С-бокс се налази узводно од Р.СR са којим је делимично преклопљен. С протеин система EcoRV има релативно висок афинитет за формирање хомодимера (Zheleznaia et al., 2003). При ниским концентрацијама С протеина, димер везан за DBS активира транскрипцију са Р.СR, док је при високим концентрацијама С димер везан за PBS инхибира. Међутим, за разлику од претходно описаних система, у овом систему се С димери везују за ДНК у потпуном одсуству кооперативности (Semenova et al., 2005). Везивање С протеина за ДНК индиректно модулише активност промотора Р.М, директним утицајем на активност Р.СR који је са њим преклопљен (Rodic, Blagojevic, Zdobnov, et al., 2017).

1.2.4 Регулација експресије Р-М система Kpn2I

Локације промотора и везивних места за С протеин у Р-М систему Kpn2I су први пут прецизно мапирани сарадници наше истраживачке групе (Klimuk et al., 2018). Генетичка организација овог система се може назвати атипичном (Слика 1D): ген за рестрикциону ендонуклеазу се преписује са слабог, конститутивно активног промотора, док су промотори гена за метилтрансферазу и гена за С протеин међусобно делимично преклопљени, супротно усмерени, и у компетицији за везивање РНК полимеразе (Klimuk et al., 2018; Lubys et al., 1999). Два места за везивање С димера која чине С-бокс су позиционирана унутар гена за С протеин, а притом се делимично преклапају са јаким промотором Р.М. Кинетика формирања конфигурације тетрамера С протеина на ДНК није детаљно изучена, али је показано да два везана димера С протеина ометају везивање РНК полимеразе за Р.М, тј. делују као репресори транскрипције гена за метилтрансферазу (Klimuk et al., 2018). Додатно, ово је први Р-М систем за који је показано деловање регулације транскрипције гена и на нивоу елонгације транскрипције, поред очекиване регулације на нивоу иницијације транскрипције (Klimuk et al., 2018). Конкретно, експериментално је показано да С протеини на ДНК „бломирају пут“ РНК полимеразе која преписује ген за С протеин (енг. *roadblock effect*). Међутим, процењена ефикасност ове блокаде је релативно ниска – око половине РНК полимеразе успе да прође кроз блокаду, а да не дисосује са ДНК, и да доврши елонгацију транскрипта. Овим је успостављена делимична негативна ауторегулација експресије С протеина. Места за везивање С протеина унутар његовог гена би могла бити широко распрострањена регулаторна карактеристика Р-М система, судећи по биоинформатичким предвиђањима (Klimuk et al., 2018).

1.3 Системи CRISPR-Cas

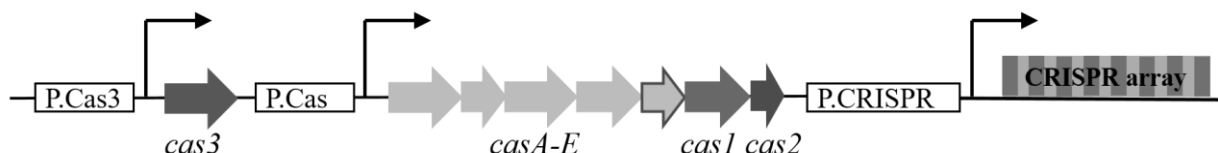
CRISPR-Cas су, у поређењу са рестрикционо-модификационим, напреднија врста прокариотских система са имунском, тј. функцијом одбране ћелије од стране ДНК (Barrangou et al., 2007; Brouns et al., 2008; Goldberg & Marraffini, 2015; Hille & Charpentier, 2016; Makarova et al., 2006; F. J. Mojica et al., 2000). Ови системи су у стању да специфично препознају и уклоне ДНК којом је ћелија била инфицирана у прошлости и чији су део секвенце узорковали и сачували, обезбеђујући дугорочни имунитет бактерије на њу (Barrangou et al., 2007; Hille & Charpentier, 2016). Основне компоненте система представљају CRISPR низ и кластер *cas* гена који се налазе у геному бактерије. CRISPR низ је локус који садржи низ истоветних директних поновака дугих ~30 бп које раздвајају тзв. спејсери (енг. *spacer*, у преводу разделник), фрагменти међусобно сличне дужине, а различитих секвенци (Barrangou et al., 2007; Brouns et al., 2008; Hille & Charpentier, 2016; Makarova et al., 2006). Анализирањем ДНК секвенци спејсера, установљено је да је велики број њих комплементаран са деловима генома различитих бактериофага, али и са плазмидима и другим мобилним генетичким елементима (Barrangou et al., 2007; Bolotin et al., 2005; Hille & Charpentier, 2016; Makarova et al., 2006; F. J. M. Mojica et al., 2005). Осим у одбрани прокариотских ћелија од стране ДНК, CRISPR-Cas системи имају улогу и у поправци оштећења ДНК, ендогеној регулацији гена и еволуцији генома (Hille & Charpentier, 2016). Од њиховог открића 1987. (Ishino et al., 1987) па до данас, пронашли су запањујући број биотехнолошких примена (Pickar-Oliver & Gersbach, 2019).

Одбрамбени механизам деловања система CRISPR-Cas обухвата три фазе које су назване адаптација, биогенеза *cr*RNA и интерференција (Hille & Charpentier, 2016). По појави стране, нпр. бактериофагне ДНК у ћелији, комплекс који се састоји од протеина Cas1 и Cas2 преузима њен фрагмент, тзв. протоспејсер, који се налази у близини PAM (енг. *protospacer adjacent motif*) секвенце, и уграђује га у геномски локус, CRISPR низ, као нови спејсер, што је познато као адаптација. У посебном процесу адаптације са прајмером (енг. *primed adaptation*), препознавање секвенце раније детектоване стране ДНК стимулише појачано сакупљање нових спејсера пореклом из те ДНК већ познатог вируса. На тај начин се узоркују и уграђују у CRISPR низ секвенце сличне већ постојећим спејсерима, што ћелији омогућава да развије резистенцију и на мутиране секвенце бактериофагних генома (Künne et al., 2016; Musharova et al., 2017; Sternberg et al., 2016). CRISPR-Cas системи су подељени на две класе унутар којих се разликује више типова и под-типова. У класи 1, разградњу препознате стране ДНК обавља мултипротеински комплекс, док у класи 2 ту улогу има само један протеин (Makarova et al., 2015). Транскрипцијом CRISPR низа се добија дугачак молекул *pre-cr*RNA из којег се, затим, процесом специфичне обраде исецају мале *cr*RNA које се састоје из појединачних спејсера оивичених деловима директних поновака (Brouns et al., 2008; Djordjevic et al., 2012; Pougach et al., 2010; Pul et al., 2010). У системима типа I из класе I, чија се регулација проучава у овој тези, за специфичну обраду *pre-cr*RNA је задужен протеин Cas6. Мултипротеински комплекс "каскада" користи *cr*RNA за препознавање ДНК одговарајућег вируса, на основу комплементарног спаривања њихових секвенци, и затим регрутује протеин Cas3 који уништава страну ДНК својом егзонуклеазном активношћу, што је познато као интерференција (Al-Attar et al., 2011; Brouns et al., 2008; Goldberg & Marraffini, 2015; Hille & Charpentier, 2016).

1.3.1 Регулација експресије система CRISPR-Cas Типа I-E

Динамика експресије система CRISPR-Cas Типа I-E, који се налази у геному *E. coli*, је остала неиспитана упркос распрострањеном коришћењу ове бактерије као експерименталног модел система, као и великом научном интересовању за системе CRISPR-Cas. Разлог томе је и даље непознат окидач за покретање активности система *in vivo* у овој бактерији, будући да је запажено да систем остаје неактиван чак и у присуству вирусне инфекције у стандардним условима раста ћелија (Patterson et al., 2017; Ratner et al., 2015). Анализа порекла спејсер секвенци CRISPR низова уверљиво указује на то да систем није у потпуности онеспособљен

током еволуције и да може имати више функција у *E. coli*, па се сматра да постоје услови у којима се он активира (Díez-Villaseñor et al., 2010). О појединим елементима и механизмима регулације експресије система Типа I-E постоје експерименталне информације или теоријска предвиђања, који потичу превасходно из истраживања спроведених на бактеријама *E. coli* и *Salmonella enterica* (Medina-Aparicio et al., 2011; F. J. M. Mojica & Díez-Villaseñor, 2010; Patterson et al., 2017; Pul et al., 2010; Richter et al., 2012; Westra et al., 2010). Типична генетичка организација CRISPR-Cas система Типа I-E из *E. coli* K12 је илустрована на Слици 2 (Brouns et al., 2008; F. J. M. Mojica & Díez-Villaseñor, 2010; Westra et al., 2010). CRISPR низ и кластер *cas* гена су под контролом два промотора који су овде означени са P.CRISPR и P.Cas, док се ген за протеин Cas3 преписује са засебног промотора.



Слика 2. Организација гена CRISPR-Cas система Типа I-E у *E. coli* K12. CRISPR низ је представљен пругастим правоугаоником чије тамније пруге одговарају директним поновцима, а светлије спејсерима. *Cas* гени су приказани широком стрелицама, док правоугаоници означени са P.Cas3, P.Cas и P.CRISPR представљају, редом, промоторе за транскрипцију *cas3* гена, *cas* оперона и CRISPR низа. Танке стрелице указују на смер транскрипције са датих промотора. Гени *casA*, *casB*, *casC*, *casD* и *casE* формирају комплекс каскада, при чему ген *casE* кодира протеин Casбе за обраду pre-crRNK. Слика адаптирана из (Rodic et al., 2018).

У стандардним условима раста културе, промотори са којих се врши транскрипција CRISPR низа и *cas* гена су утишани низовима везаних H-NS и LRP протеина (Pougach et al., 2010; Pul et al., 2010; Westra et al., 2010), глобалних регулатора који се у великом броју везују за ДНК са високом кооперативношћу (Bouffartigues et al., 2007). Поред ова два, показано је да у репресији промотора *cas* оперона може да учествује и протеин CRP (Yang et al., 2014). Није јасно под којим условима се ова репресија уклања, али је показано да се ово може постићи експресијом високе концентрације глобалних регулатора као што је протеин LeuO (Medina-Aparicio et al., 2011; Pul et al., 2010; Westra et al., 2010). Ниво овог и других глобалних активатора транскрипције, међу којима су LexA и BaeR-S, се може повећати у ћелији у одговору на стрес мембране који се може јавити, између осталог, као последица напада бактериофага, па је предложено да би стрес мембране могао бити окидач за активацију експресије гена система CRISPR-Cas I-E (Patterson et al., 2017; Ratner et al., 2015; Richter et al., 2012). Осим што активира транскрипцију са промотора система CRISPR-Cas, LeuO може и да индиректно утиша експресију сопственог гена (Chen et al., 2001; Rodic et al., 2018; Stratmann et al., 2012).

Регулација система је проучавана и на нивоу обраде CRISPR транскрипта (Brouns et al., 2008; Djordjevic et al., 2012; Pougach et al., 2010). Примећено је да висока експресија протеина Casбе у *E. coli* доводи до релативно малог смањења већ ниског равнотежног нивоа pre-crRNK (<10 молекула по ћелији), на рачун којег се равнотежни ниво crRNK повећава за ~2 реда величине (на ~1000 молекула по ћелији) (Pougach et al., 2010; Pul et al., 2010). Наша група је моделовала процес настанка crRNK од pre-crRNK и показала да само пораст концентрације протеина Casбе задуженог за обраду није довољан да објасни ову промену у равнотежним нивоима молекула (Djordjevic et al., 2012). Заправо, брза неспецифична деградација pre-crRNK неидентификованом нуклеазом је истакнута као један од веома утицајних регулаторних својстава система. Одвлачење pre-crRNK са пута неспецифичне разградње на пут специфичног сечења којим настају crRNK објашњава мало уочено смањење количине pre-crRNK у ћелији по активацији обраде (Djordjevic et al., 2012; Pougach et al., 2010; Pul et al., 2010).

1.4 Принципи термодинамичког моделовања регулације транскрипције

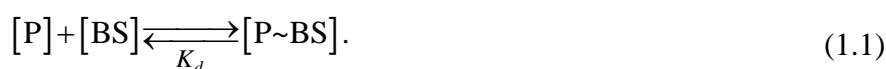
Процедура моделовања регулације транскрипције заснована у (Shea & Ackers, 1985) и формализована у (Buchler et al., 2003) обухвата неколико корака. Прво је потребно идентификовати регион ДНК који садржи промотор и остале елементе релевантне за његову транскрипциону активност, и утврдити у каквим се све стањима, тј. конфигурацијама тај фрагмент може наћи. Конфигурације које се подразумевају су слободна ДНК, за коју није везан ниједан регулатор транскрипције, и конфигурација у којој је промотор заузет РНК полимеразом у одсуству других учесника, када је његова транскрипциона активност базална. Присуство других конфигурација, којима се могу постићи активација и репресија промотора, зависи од транскрипционих регулатора под чијом је контролом дати промотор, и конкретних образаца и кинетике њиховог комбинаторног везивања за ДНК (Buchler et al., 2003). На пример, у околини промотора *lac* оперона у *E. coli*, који кодира ензиме за метаболизам лактозе, налазе се места за везивање протеина CAP, који активира транскрипцију, и Lac репресора (Phillips et al., 2012). Како се они за ДНК везују независно један од другог, могу се очекивати конфигурације ДНК фрагмента за коју је везан само Lac репресор и фрагмента за који је осим Lac репресора везан и активатор CAP. CAP веома ефикасно регрутује РНК полимеразу на промотор, па се у одсуству везаног Lac репресора може формирати конфигурација CAP и РНК полимеразе на ДНК, чим се CAP веже за своје везивно место.

У следећем кораку се одређују статистичке тежине конфигурација, тј. релативне вероватноће њиховог налажења у ћелији, према теорији статистичке термодинамике (Shea & Ackers, 1985). Притом се уводи важна претпоставка да се кинетика везивања протеина за ДНК може представити равнотежним хемијским реакцијама. Ова апроксимација је оправдана јер се процес експресије гена, који се прати на излазу модела, одиграва на знатно споријој временској скали од процеса везивања и дисоцијације протеина са ДНК (Ingalls, 2013). Статистичке тежине конфигурација се изражавају у функцији концентрација протеина који их граде и јачина протеин-ДНК и протеин-протеин интеракција, тако да се за познате вредности ових величина, помоћу модела може предвидети интензитет транскрипционе активности промотора (Dresch et al., 2013; Rodic et al., 2018).

Транскрипциона активност промотора, мерена концентрацијом синтетисаних транскрипата у јединици времена, се изражава у функцији статистичких тежина његових конфигурација, при чему се сматра директно пропорционалном вероватноћи формирања транскрипционо активних конфигурација, тј. оних конфигурација у којима је за промотор везана РНК полимераз (Shea & Ackers, 1985). Овим се стопа експресије гена имплицитно изједначава са стопом иницијације транскрипције, тј. занемарују се сви низводни догађаји који потенцијално могу довести до прекидања синтезе транскрипта, али се неки ефекти, као што је ефикасност елонгације транскрипта, могу укључити у модел (Klimuk et al., 2018). Овакви једноставни модели транскрипције гена прокариота имају веома добру моћ предвиђања (Ay & Arnosti, 2011; Bintu et al., 2005; Buchler et al., 2003; Hawley & McClure, 1982; Morrison et al., 2021; Vilar & Leibler, 2003), а успешно су примењени и на експресију гена еукариота (Ay & Arnosti, 2011; Gertz et al., 2009; He et al., 2010; Segal et al., 2008).

1.4.1 Извођење статистичких тежина полазећи од равнотежних биохемијских реакција

Следећа равнотежна биохемијска реакција описује везивање протеина P за своје везивно место BS (од енг. *Binding Site*) на ДНК, при чему настаје комплекс P~BS (Rodic et al., 2018; Snerpen & Zocchi, 2005):



Кинетика ове реакције зависи од концентрација учесника и вредности равнотежне константе дисоцијације,

$$K_d = ([P][BS])/[P\sim BS]. \quad (1.2)$$

Фрагмент ДНК са везивним местом BS се може наћи у две конфигурације: са и без везаног протеина. Укупна концентрација посматраних фрагмената у ћелији представља збир концентрација везивних места у ове две конфигурације, $[BS]_{\text{tot}} = [P\sim BS] + [BS]$. Вероватноћа налажења конфигурације са везаним протеином је $P_{ON} = [P\sim BS]/[BS]_{\text{tot}}$, док је вероватноћа конфигурације слободне ДНК, $P_{OFF} = [BS]/[BS]_{\text{tot}}$. Статистичка тежина везаног протеина за ДНК је дата релативном вероватноћом налажења заузетог BS у односу на вероватноћу слободног BS и, узимајући у обзир једначину (1.2), износи (Rodic et al., 2018; Sneppen & Zocchi, 2005):

$$\frac{P_{ON}}{P_{OFF}} = \frac{[P\sim BS]}{[BS]} = \frac{[P]}{K_d}. \quad (1.3)$$

1.4.2 Извођење статистичких тежина према теорији статистичке термодинамике

Ћелија са молекулима протеина који се крећу слободном дифузијом и фрагментом ДНК на коме се налази везивно место за протеин се, према теорији статистичке термодинамике, може представити затвореним системом (s) у контакту са „резервоаром“ (r) са којим размењује енергију (Phillips et al., 2012; Rodic et al., 2018). Ако је резервоар много већи од система, енергија *укупног* система који чине ћелија (енергије E^s) и резервоар (енергије E^r),

$$E = E^s + E^r, \quad (1.4)$$

се може сматрати приближно константном докле год важи да је $E^s \ll E$, а укупан систем изолован. У изолованом систему, према Другом закону термодинамике, расподела енергије између система и резервоара ће се мењати док се не успостави она при којој је укупном систему доступан највећи број микростања, тј. јединствених просторних распореда честица система (Phillips et al., 2012). У случају замишљене ћелије, различити распореди молекула протеина представљају различита микростања система. Број микростања, Ω , укупног изолованог система у коме систему припада одређена енергија E_i^s је дат производом броја микростања у којима може да се нађе систем (Ω^s) и броја микростања доступних резервоару (Ω^r) у тим условима:

$$\Omega(E, E_i^s) = \Omega^s(E_i^s) \cdot \Omega^r(E - E_i^s). \quad (1.5)$$

Према дефиницији ентропије, $S = k_B \ln(\Omega)$, број микростања се може изразити и на следећи начин:

$$\Omega = \exp(S/k_B), \quad (1.6)$$

где је k_B Болцманова константа. Како изоловани систем тежи расподели енергије са максималним бројем њему доступних микростања, вероватноћа да систему припадне енергија E_i^s је директно пропорционална $\Omega(E, E_i^s)$, односно кад се искористе једначине ((1.5)-(1.6)):

$$\begin{aligned}
P(E_i^s) &\sim \Omega^s(E_i^s) \cdot \exp\left(S^r(E - E_i^s)/k_B\right) \\
&\approx \Omega^s(E_i^s) \cdot \exp\left(\left(S^r(E) - \frac{dS^r}{dE} \cdot E_i^s\right)/k_B\right) \\
&\sim \Omega^s(E_i^s) \cdot \exp\left(-E_i^s/(k_B \cdot T)\right)
\end{aligned} \tag{1.7}$$

при чему је у другом реду једначине (1.7) зависност ентропије резервоара од енергије линеаризована у околини $S^r(E)$, што добро апроксимира оригиналну зависност за $E^s \ll E$ (Rodic et al., 2018; Sneppen & Zocchi, 2005). У трећем реду једначине (1.7) је искоришћена термодинамичка дефиниција температуре, односно да инверзна температура одговара парцијалном изводу ентропије по енергији, при константној запремини и броју честица:

$$\left(\frac{\partial S}{\partial E}\right)_{V,N} = \frac{1}{T}, \tag{1.8}$$

при чему је константа $\exp(S^r(E))$ изостављена. Дакле, вероватноћа да систем има енергију E_i^s је сразмерна производу броја његових микростања (први члан у трећем реду једначине (1.7)) и тзв. Болцмановог фактора (други члан у трећем реду једначине (1.7)), који представља статистичку тежину појединачног микростања система, тј. ненормализовану вероватноћу успостављања једног одређеног микростања система при енергији E_i^s (Sneppen & Zocchi, 2005). Притом, израз у трећем реду једначине (1.7) управо одговара статистичкој тежини макростања система енергије E_i^s , тј. стања које обједињује сва микростања која се могу успоставити при овој енергији. На пример, микростања ћелије са молекулима протеина и једним ДНК везивним местом се могу сврстати у две групе, тј. макростања: 1) макростање у коме је везивно место слободно, а сви протеини су у раствору и 2) макростање у коме је један протеин везан за ДНК док су остали у раствору (Phillips et al., 2012). Да би се одредила вероватноћа макростања система одређене i -те енергије E_i^s , његову статистичку тежину је потребно поделити збиром статистичких тежина свих макростања у којима се систем може наћи:

$$P(E_i^s) = \frac{\Omega^s(E_i^s) \cdot e^{-E_i^s/(k_B \cdot T)}}{\sum_i \left(\Omega^s(E_i^s) \cdot e^{-E_i^s/(k_B \cdot T)} \right)}. \tag{1.9}$$

Статистичке тежине два макростања ћелије са везивним местом за протеин, како је објашњено, треба да се изразе као производ два члана – броја микростања система и Болцмановог фактора (Rodic et al., 2018). Болцманов фактор зависи од енергије система коју ћемо означити са ε_{bound}^{bs} за макростање са заузетим, односно са $\varepsilon_{unbound}^{bs}$ за макростање са слободним везивним местом. Број микростања система у овом примеру одговара броју начина на које слободни протеини могу да се распореде у шестодимензионалном фазном простору ћелије. Наиме, молекула протеина са кинетичком енергијом $\varepsilon_k = p^2/(2m)$ (где је p импулс, а m маса протеина) који се слободно креће по унутрашњости ћелије се може окарактерисати са три димензије његове позиције у простору (r) и три димензије импулса (p). Према принципу неодређености у квантној механици, овакав шестодимензионални фазни простор може да се издели на дискретне јединице величине Планкове константе (h), при чему у једну јединицу може да се смести само један протеин (Sneppen & Zocchi, 2005; Stowe, 2007). У случају слободног везивног места, свих N молекула протеина се слободно креће, па статистичка тежина овог макростања, Z_{OFF} , зависи од броја различитих распореда N истих честица у фазном простору система и енергије система $\varepsilon_{unbound}^{bs}$:

$$Z_{OFF} = \frac{1}{N!} \left(\int_V \int \frac{d^3 r \cdot d^3 p}{h^3} e^{-p^2/(2mk_B T)} \right)^N e^{-\varepsilon_{unbound}^{bs}/(k_B T)}, \quad (1.10)$$

$$\propto k^N \rho^{-N} e^{-\varepsilon_{unbound}^{bs}/(k_B T)}$$

док у израз за статистичку тежину макростања са једним протеином везаним за ДНК, Z_{ON} , улазе број распореда преосталих $N - 1$ протеина и енергија ε_{bound}^{bs} :

$$Z_{ON} = \frac{1}{(N-1)!} \left(\int_V \int \frac{d^3 r \cdot d^3 p}{h^3} e^{-p^2/(2mk_B T)} \right)^{N-1} e^{-\varepsilon_{bound}^{bs}/(k_B T)}, \quad (1.11)$$

$$\propto k^{N-1} \rho^{-(N-1)} e^{-\varepsilon_{bound}^{bs}/(k_B T)}$$

где је $k = (2mk_B T \pi / h^2)^{3/2}$, а $\rho = N/V$ (V је запремина ћелије) (Phillips et al., 2012; Rodic et al., 2018; Sneppen & Zocchi, 2005). Ако је $Z = Z_{ON} + Z_{OFF}$ збир статистичких тежина свих могућих макростања система (а овај систем их има укупно два), однос вероватноћа да везивно место на ДНК буде заузето, односно слободно, је дат следећим изразом који користи J -не (1.10)-(1.11):

$$\frac{P_{ON}}{P_{OFF}} = \frac{Z_{ON}}{Z} \cdot \left(\frac{Z_{OFF}}{Z} \right)^{-1} = \frac{\rho}{k} e^{-\Delta\varepsilon/(k_B T)}, \quad (1.12)$$

где је $\Delta\varepsilon = \varepsilon_{bound}^{bs} - \varepsilon_{unbound}^{bs}$ енергија везивања протеина за ДНК. Како процес везивања подразумева и конформационе промене молекула, енергија везивања је дата Гибсовом (енг. Gibbs) слободном енергијом, ΔG (Sneppen & Zocchi, 2005). Овај израз (J -на (1.12)) је еквивалентан једначини (1.3) изведеној полазећи од система равнотежних биохемијских реакција, при чему је $[P] = \rho$, а $K_d = k \cdot \exp(\Delta\varepsilon/(k_B T))$.

1.5 Принципи динамичког моделовања експресије протеина

Динамички модели имају значајно место у испитивању експресије генских мрежа (Aguda et al., 2008; Alon, 2006; Ay & Arnosti, 2011; Ingalls, 2013; Le Novère, 2015; Williams et al., 2013). Понашање мреже регулаторних тј. биохемијских реакција у времену се може описати системом обичних диференцијалних једначина (Ay & Arnosti, 2011; Ingalls, 2013). Како концентрација дате хемијске врсте – компоненте мреже – зависи од процеса у којима она настаје и којима се елиминише из система, брзина њене промене са временом је дата резултантним збиром брзина релевантних процеса. Ти процеси могу бити представљени различитим реакцијама, од елементарних на које се примењује закон о дејству маса, па до сложених биохемијских (нпр. Михаелис-Ментенов кинетички закон), али оно што је заједничко у изразима за њихове брзине је зависност од концентрација других молекулских врста, реактаната (Ingalls, 2013; Rodic et al., 2018).

Динамички модел експресије гена се састоји из обичних диференцијалних једначина које описују промене концентрација транскрипата и протеина посматраног система (Rodic et al., 2018). За сваки i -ти ген система се може написати пар диференцијалних једначина – једна за промену концентрације транскрипта ($mRNA_i$), тј. информационе (енг. messenger) РНК настале транскрипцијом гена, и једна за промену концентрације протеина кодираног геном ($protein_i$):

$$\frac{dmRNA_i(t)}{dt} = \varphi_i(t) - \lambda_i^{mRNA} \cdot mRNA_i(t) \quad (1.13)$$

$$\frac{dprotein_i(t)}{dt} = \kappa_i \cdot mRNA_i(t) - \lambda_i^{protein} \cdot protein_i(t) \quad (1.14)$$

Први чланови једначина (1.13) и (1.14) представљају брзине процеса у којима се транскрипт и протеин синтетишу: транскрипциону активност промотора датог гена (φ_i), односно брзину синтезе протеина транслацијом са константом κ_i . Транскрипциона активност промотора може представљати пример брзине комплексног процеса регулације промотора који се може моделовати коришћењем приступа статистичке термодинамике (видети поглавље 1.4). Други чланови једначина (1.13) и (1.14) представљају брзине процеса у којима се транскрипт и протеин елиминишу из система са константама λ_i^{mRNA} , односно $\lambda_i^{protein}$. Ови процеси типично подразумевају разградњу молекула, али и ефективно разблаживање њихових концентрација приликом деобе ћелије (Klumpp et al., 2009; Morozova et al., 2016).

Систем нелинеарних диференцијалних једначина се најчешће не може решити аналитички, већ се различите нумеричке методе користе за одређивање апроксимираног решења (Au & Arnosti, 2011). Моделовање система обичним диференцијалним једначинама укључује претпоставку о детерминистичкој природи процеса, тј. да је концентрација одређене хемијске врсте у наредном временском тренутку једнозначно одређена тренутним концентрацијама реактанта и скупом кинетичких правила која описују процесе у којима учествује. Ова претпоставка је валидна у константним физичко-хемијским условима средине, при чему су молекули хомогено распоређени у простору и присутни у релативно великим бројевима, тако да се њихове концентрације могу сматрати континуалним варијаблама (Ingalls, 2013). Ако концентрације молекула варирају и у времену и у простору, за моделовање система је потребно користити парцијалне диференцијалне једначине. Додатно, у случају малог броја молекула у систему, где случајни процеси могу имати незанемарљив утицај на њихову експресију током времена, уводе се стохастички модели.

1.5.1 Експериментално одређивање динамичке криве експресије протеина Р-М система у појединачној ћелији

Одређивање криве промене концентрације одређеног протеина током времена у појединачној ћелији путем експерименталног мерења је захтеван задатак, па је динамика успостављања Р-М система Типа II у бактеријској ћелији регистрована у само два случаја (Morozova et al., 2016; Mruk & Blumenthal, 2008; Rodic et al., 2018). Уколико се концентрација протеина експериментално прати на нивоу целе популације ћелија (енг. *bulk assay*), да би се тачно реконструисала динамика на нивоу једне ћелије неопходна је синхронизована култура ћелија. То је успешно постигнуто скоро истовременим уносом Р-М система PvuII у све ћелије културе коришћењем бактериофага M13 као вектора (Mruk & Blumenthal, 2008). Како су се од тог момента ћелије врло синхронизовано делиле, било је могуће одредити динамику протеина на нивоу једне ћелије. Током времена су праћени нивои транскрипата и ензимске активности рестрикционе ендонуклеазе и метилтрансферазе, да би затим вредности добијене за целу популацију биле подељене бројем ћелија у њој.

У другом случају, интензитет флуоресценције пропорционалан концентрацији протеина Р-М система Esp1396I је мерен у експерименту на нивоу појединачних ћелија (енг. *single-cell*), у култури пореклом од једне једине ћелије трансформисане плазмидом као носачем гена система (Morozova et al., 2016). У сврху овог експеримента, конструисан је модификовани Р-М систем Esp1396I чији су гени за рестрикциону ендонуклеазу и метилтрансферазу фузионисани са генима флуоресцентних протеина *mCherry* и *Venus*, редом. Показано је да је систем притом остао функционалан, односно да додати флуоресцентни протеини нису довели до губитка

функције ензима. Интензитет флуоресценције оба протеина и раст културе су праћени флуоресцентном микроскопијом.

Из ових примера је јасно да би експериментално испитивање значаја различитих регулаторних својстава Р-М система за динамику његовог успостављања у бактеријској ћелији било веома захтевно и изазовно. Генска кола система са уведеним различитим изменама његових регулаторних својстава би морала да буду конструисана, а затим би требало пратити динамику експресије транскрипата или протеина сваког од измењених система на нивоу ћелије. Симулација оваквих експеримената *in silico* је далеко лакше изводљива и нуди могућност фокусирања циљева будућих експеримената на аспекте истакнуте предвиђањима модела.

2 Хипотезе и циљеви истраживања

Општи циљ истраживања представљеног у овој докторској дисертацији је проучавање динамике експресије гена бактеријских имунских система од тренутка њихове активације у ћелији, и то из угла захтева које намећу основне функције ових система. Рестрикционо-модификациони и CRISPR-Cas системи деле функцију експресије потенцијално штетних молекула у ћелији – рестрикционе ендонуклеазе, односно crRNK - који могу угрозити интегритет генома домаћина (Djordjevic, 2013; Rodic et al., 2018). Да би рестрикциона ендонуклеаза P-M система била корисна за бактеријску ћелију, уклањањем вирусних и осталих страних молекула ДНК, њена експресија мора пратити строго контролисан временски образац, што је предложено у литератури, али не и детаљно испитано. Конкретно, издвојиле су се две особине динамике успостављања P-M система у новонастањеној бактеријској ћелији – одложена синтеза рестрикционе ендонуклеазе у односу на акумулацију метилтрансферазе и опадање нивоа метилтрансферазе у каснијем интервалу – као кључни услови за безбедно, а ефикасно функционисање система. Међутим, динамика експресије P-M система је експериментално мерена у само два случаја у којима су теоријска очекивања делимично потврђена. Са друге стране, динамику активације система CRISPR-Cas Типа I-E је још увек немогуће експериментално регистровати јер сигнал који би довео до покретања утишаног система у *E. coli* није откривен.

Стога је, у овој дисертацији, динамика бактеријских имунских система предвиђена и анализирана теоријски, конструисањем математичких модела регулације експресије гена система, заснованих на доступним експерименталним информацијама. Како је полазни циљ био уочавање основних, квалитативних принципа динамике, коришћени су детерминистички динамички модели са термодинамички моделованом регулацијом транскрипције. Овим путем је анализирана динамика успостављања четири P-M система, при чему је динамика система EcoRV (Rodic, Blagojevic, Zdobnov, et al., 2017) и Kpn2I (Klimuk et al., 2018) по први пут предвиђена. Модел регулације система AhdI, претходно постављен од стране наше групе, је адаптиран ради анализе квалитативних особина динамике овог система (Rodic, Blagojevic, Zdobnov, et al., 2017), а модел система Esp1396I је конструисан да би се његова теоријска предвиђања упоредила са подацима добијеним мерењем динамике протеина овог система на нивоу појединачних ћелија (Morozova et al., 2016). Непозната динамика активације система CRISPR-Cas је теоријски предвиђена на основу сазнања о општим механизмима регулације описаним за тај систем (Rodic, Blagojevic, Djordjevic, et al., 2017).

Узимајући у обзир специфичне захтеве имунске функције P-M система – ефикасно и безбедно успостављање система по уласку у бактеријску ћелију без система – формулисали смо хипотезу о универзалним динамичким принципима дизајна регулације ових система (Rodic, Blagojevic, Zdobnov, et al., 2017). Конкретно, предложили смо које би три особине динамике успостављања P-M система у новом домаћину могле представљати ограничења која су усмерила њихову еволуцију према данас присутним архитектурама и механизмима регулације. Прво, као што је већ широко прихваћено у литератури, неопходно је остварити почетно кашњење у експресији рестрикционе ендонуклеазе за експресијом метилтрансферазе, како активност рестрикционе ендонуклеазе не би била усмерена на незаштићене циљне секвенце у геному бактерије. Друго, прелазак експресије рестрикционе ендонуклеазе са веома ниског на висок ниво, након почетног кашњења, би требало да се заврши у што краћем временском периоду, како би ћелија што брже добила ефикасну заштиту од стране ДНК. Трећи и последњи предложени динамички принцип је одржавање стационарног стања без великих флукуација у концентрацијама ензима услед којих би оптималан баланс активности

рестрикционе ендонуклеазе и метилтрансферазе могао у неком тренутку да буде нарушен, што је опасно по ћелију (Djordjevic, 2013). Према овој хипотези, различите комбинације регулаторних елемената и својстава које се срећу у природним Р-М системима су састављене, односно задржане, јер обезбеђују динамику експресије система са предложеним особинама. Услед тога, пертурбисање кључних регулаторних својстава неког Р-М система би требало да резултира ремећењем барем једног од предложена три динамичка принципа. Слично, увођење регулаторног својства које у једном систему има веома значајну улогу у креирању жељених динамичких особина, у други Р-М систем, са другачијом комбинацијом регулаторних својстава, би требало да измени динамику експресије тог система у смеру губитка жељених особина. Како бисмо могли да измеримо усаглашеност система са динамичким принципима приликом пертурбисања својстава његове регулације, дефинисали смо квантитативне величине које описују степен развијености предложене три особине динамике експресије.

Хипотезу смо тестирали анализирањем два Р-М система – AhdI и EcoRV (Rodic, Blagojevic, Zdobnov, et al., 2017). Адаптирали смо раније постављен модел регулације система AhdI са конвергентном архитектуром и пертурбисали смо његова три карактеристична регулаторна својства: 1) ниског афинитета С протеина за формирање хомодимера, услед којег се у раствору налазе претежно мономери, 2) ниже ефикасности транслације цистрона за С протеин у поређењу са цистронима осталих протеина система која је последица одсуства наменског места за везивање рибозома, и 3) високе кооперативности у везивању два С димера за С-бокс. Притом смо пратили како се динамика експресије мења у односу на предвиђену динамику *wt* система, односно да ли се жељене особине динамике ремете. Пертурбисањем истих регулаторних својстава у Р-М систему EcoRV смо ефективно увели у овај систем карактеристична својства система AhdI. Наиме, EcoRV се, за разлику од система AhdI, одликује дивергентном архитектуром, али и супротним наведеним регулаторним својствима: 1) С протеини у раствору претежно формирају димере, 2) цистрон за С протеин садржи место за везивање рибозома, па је транслација цистрона сва три протеина приближно исте ефикасности, и 3) С димери се за С-бокс везују у потпуном одсуству кооперативности. Притом смо, као и приликом симулације динамике система AhdI, одређивали вредности квантитативних величина за мерење заступљености динамичких особина, како бисмо утврдили у ком смеру се ове особине мењају пертурбисањем кључних регулаторних својстава система.

Осим за систем EcoRV, у овој докторској дисертацији је дато прво предвиђање динамике експресије и за систем Kpn2I (Klimuk et al., 2018). Наши сарадници су извели низ експеримената којима су испитали механизме регулације гена и притом показали да С протеини везани за ДНК у овом систему регулишу транскрипцију гена за метилтрансферазу на нивоу иницијације, али и сопственог гена на нивоу елонгације транскрипције. Наш циљ је био да поставимо модел регулације система Kpn2I који описује све уочене механизме и калибришемо га користећи резултате експеримената наших сарадника. Симулацијом динамике смо желели да проверимо да ли регулација овог система може да обезбеди почетну акумулацију метилтрансферазе, односно одлагање активности рестрикционе ендонуклеазе који се сматрају неопходним приликом успостављања система. Такође, пертурбисали смо ефикасност везаних С протеина у блокирању пута РНК полимеразе приликом елонгације транскрипције гена за С протеин варирајући њену вредност између 0 и 100% и притом смо симулирали динамику система. Полазна хипотеза у основи ове анализе је била да је експериментално показана ниска ефикасност блокаде (<50%) у *wt* систему резултат финог подешавања вредности овог параметра током еволуције система, тј. да је фиксирана вредност релевантна за оптималну динамику експресије ензима система. У складу са тим, очекује се да промена вредности ефикасности блокаде наруши оптималну динамику експресије, односно резултира динамиком која би могла бити кобна по домаћина или неодговарајућа за обављање имунске функције система.

Модел регулације система Esp1396I смо конструисали са циљем да његова предвиђања упоредимо са експерименталним подацима наших сарадника који су динамику промене нивоа рестрикционе ендонуклеазе и метилтрансферазе овог система мерили на нивоу појединачних ћелија (Mogozova et al., 2016). Стопа транскрипције оперона са генима за C протеин и рестрикциону ендонуклеазу у систему Esp1396I је регулисана истим механизмом који је детаљно проучен на систему AhdI. Експериментално потврђен модел регулације транскрипције овим механизмом, који подразумева везивање два C димера за C-box који учествују, прво у активацији, а затим и у репресији промотора, је искоришћен у моделовању регулације Esp1396I. Карактеристична својства система Esp1396I која су такође описана моделом су репресија промотора гена за метилтрансферазу C протеином, као и транслационо спрезање цистрона за C протеин и рестрикциону ендонуклеазу. Додатно, током експеримената праћења динамике експресије система је забележена и промена стопе раста културе, па је и то унето у модел, као временска зависност стопе уклањања молекула из ћелије услед њиховог разблаживања ћелијском деобом. Осталим параметрима су додељене константне вредности и поставило се питање да ли овај минимални модел, који узима у обзир позната својства и механизме система, може да репродукује експериментално регистровану динамику експресије ензима. Такође, да ли се описаном регулацијом може обезбедити жељено првенство метилтрансферазе у експресији у новој ћелији, односно кашњење експресије рестрикционе ендонуклеазе? Било каква изражена, квалитативна одступања предвиђања модела од експерименталних података би указала на регулаторни механизам или ефекат са значајном улогом који није узет у обзир и описан моделом.

Уочили смо одређене сличности између претпостављеног механизма регулације промотора у систему CRISPR-Cas Типа I-E и солидно проученог механизма регулације транскрипције C протеином, описаног на P-M системима AhdI и Esp1396I (Rodic, Blagojevic, Djordjevic, et al., 2017). Конкретно, промотори система CRISPR-Cas су утишани веома кооперативним везивањем репресора (H-NS и др.), док промотор P.CR у наведеним P-M системима утишава веома кооперативно везивање два димера C протеина. Такође, репресију у систему CRISPR-Cas може уклонити LeuO (и други активатори) који притом негативно регулише експресију свог гена (Chen et al., 2001; Stratmann et al., 2012), а слично се односи на C протеин P-M система. С обзиром на сличности у деловању регулаторних компоненти, претпоставили смо да би динамику активације система CRISPR-Cas могле да одликују квалитативне особине динамике експресије коју производи описани механизам P-M система. У том случају, непознати механизам регулације транскрипције у систему CRISPR-Cas се, у моделовању регулације овог система, може премостити његовим замењивањем сличним механизмом P-M система. Стављањем два промотора система CRISPR-Cas у моделу под контролу C протеина из система AhdI, обезбедили смо да њихове транскрипционе активности прате динамику сличну оној која би се вероватно испољила приликом природне активације система CRISPR-Cas у ћелији *E. coli*. Повезивањем модела регулације транскрипције са моделом обраде pre-crRNK, предвидели смо квалитативне карактеристике динамике активације система CRISPR-Cas, односно експресије crRNK у ћелији. Затим смо искористили модел да пертурбисањем кључних регулаторних својстава овог система и конструисањем алтернативних, редукованих механизма регулације транскрипције испитамо њихову улогу у обезбеђивању појединих динамичких особина експресије. Иако се P-M и CRISPR-Cas системи значајно разликују по архитектури и механизмима деловања, претпоставили смо да би се у њиховој динамици експресије морале уочити и неке заједничке, универзалне особине, будући да деле исту функцију одбране ћелије од стране ДНК, као и безбедне експресије потенцијално токсичних молекула. Овакве универзалне принципе смо покушали да расветлимо анализирањем предвиђене динамике система CRISPR-Cas и његових редукованих варијанти.

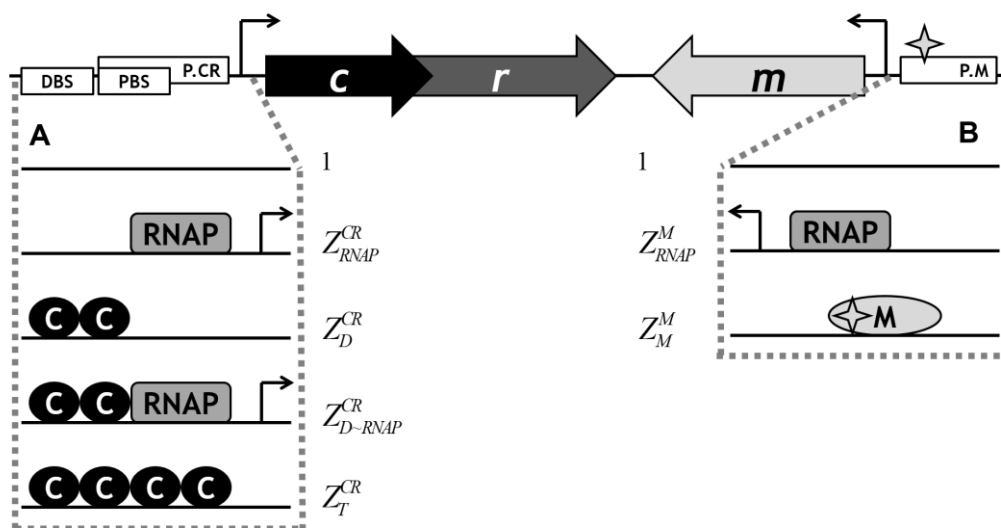
3 Методе

3.1 Биофизичко моделовање P-M система AhdI

3.1.1 Термодинамичко моделовање регулације транскрипције гена система AhdI

Претпоставка модела је да се P.CR може наћи у следећим конфигурацијама (Слика 3А) (Bogdanova et al., 2008; Rodic, Blagojevic, Zdobnov, et al., 2017; Rodic et al., 2018):

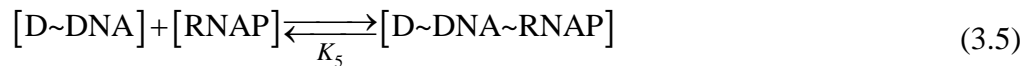
- 1) Конфигурација празног промотора, као и регулаторних секвенци у његовој околини;
- 2) Конфигурација која настаје самосталним везивањем РНК полимеразе за промотор чиме се омогућава базална транскрипција гена;
- 3) Конфигурација у којој су два димера С протеина везана за ДНК, при чему је С димер на PBS регрутован од стране С димера везаног за DBS. У овој конфигурацији промотор је транскрипционо неактиван, односно С протеин има улогу репресора;
- 4) Конфигурација у којој је РНК полимеразе регрутована на промотор од стране С димера везаног за DBS који делује као активатор транскрипције;
- 5) Конфигурација у којој је само један С димер на ДНК, везан за дистално везивно место. Услед експериментално демонстриране велике кооперативности у везивању другог С димера када је један већ везан за DBS, односно РНК полимеразе што је потврђено моделовањем, вероватноћа налажења ове конфигурације у поређењу са другим је веома ниска, због чега она није била укључена у оригинални модел (Bogdanova et al., 2008). Међутим, она се не може занемарити у условима ниске кооперативности у везивању С протеина, па је, ради испитивања ефекта пертурбисања ове карактеристике система, овде уведена (Rodic, Blagojevic, Zdobnov, et al., 2017). Конфигурација у којој је само С димер везан за PBS се занемарује због веома ниског афинитета проксималног везивног места, односно неопходности кооперативне интеракције са другим С димером.



Слика 3. Графички приказ регулације транскрипције у P-M систему AhdI. Организација гена система на плазмиду је приказана одговарајуће именованим широким стрелицама, док правоугаоници означени са P.CR P.M, DBS и PBS представљају, редом, промоторе за транскрипцију *cr* оперона и *m* гена и места за везивање С димера (дистално и проксимално). Танке стрелице указују на смер транскрипције са датих промотора. Звездича симболизује присуство циљне секвенце у региону P.M коју метилтрансфераза модификује. Дозвољене конфигурације РНК полимеразе (правоугаоник RNAP), С протеина (круг обележен са С представља мономер) и метилтрансферазе

(овал М) на ДНК су приказане за (А) Р.СР и (В) Р.М промотор. Конфигурације које садрже стрелицу су транскрипционо активне. Поред сваке конфигурације је ознака њене статистичке тежине. Слика адаптирана из (Rodic et al., 2018).

Претпостављене конфигурације Р.СР настају као резултат следећих равнотежних реакција везивања:



у којима су концентрације С протеина у форми мономера и димера означене са [Mon] и [D], концентрација слободне РНК полимеразе са [RNAP], док [DNA], [RNAP~DNA], [D~DNA], [T~DNA] и [D~DNA~RNAP] означавају концентрације, редом, слободне ДНК промоторског региона, РНК полимеразе везане за ДНК, С димера везаног за DBS, комплекса С тетрамера (два С димера) са ДНК, и комплекса који чине ДНК, С димер на DBS и РНК полимеразу на промотору. Равнотежне константе дисоцијације (K_1 - K_5) реакција (3.1)-(3.5):

$$K_1 = [\text{Mon}]^2 / [\text{D}] = k e^{\Delta G_D} \quad (3.6)$$

$$K_2 = ([\text{RNAP}][\text{DNA}]) / [\text{RNAP} \sim \text{DNA}] = k e^{\Delta G_{P.CR}} \quad (3.7)$$

$$K_3 = ([\text{D}][\text{DNA}]) / [\text{D} \sim \text{DNA}] = k e^{\Delta G_{DBS}} \quad (3.8)$$

$$K_4 = ([\text{D}][\text{D} \sim \text{DNA}]) / [\text{T} \sim \text{DNA}] = k e^{\Delta G_{PBS} + \Delta G_T} \quad (3.9)$$

$$K_5 = ([\text{D} \sim \text{DNA}][\text{RNAP}]) / [\text{D} \sim \text{DNA} \sim \text{RNAP}] = k e^{\Delta G_{P.CR} + \Delta G_{D-RNAP}} \quad (3.10)$$

се могу изразити преко одговарајућих слободних енергија интеракције у јединицама $k_B T$ (k_B – Болцманова константа, T – температура): ΔG_D – два С мономера, $\Delta G_{P.CR}$ – РНК полимеразе са промотором Р.СР, ΔG_{DBS} – С димера са DBS, ΔG_{PBS} – С димера са PBS, ΔG_T – С димера везаних за ДНК, и ΔG_{D-RNAP} – С димера везаног за DBS и РНК полимеразе везане за промотор, где је k константа пропорционалности у јединицама концентрације.

Транскрипциона активност Р.СР се, према класичној претпоставци (Shea & Ackers, 1985), сматра директно пропорционалном равнотежној вероватноћи налажења РНК полимеразе на промотору, која се изражава као удео статистичких тежина транскрипционо активних конфигурација (тј. оних које подразумевају везану РНК полимеразу) у збиру статистичких тежина свих дозвољених конфигурација промотора:

$$\varphi_{cr} = \alpha \frac{Z_{RNAP}^{CR} + Z_{D \sim RNAP}^{CR}}{1 + Z_{RNAP}^{CR} + Z_D^{CR} + Z_{D \sim RNAP}^{CR} + Z_T^{CR}} \quad (3.11)$$

У једначини (3.11), α је константа пропорционалности са јединицама концентрације cr транскрипта по јединици времена, а статистичке тежине (Z) конфигурација промотора су означене као на Слици 3 и добијају се према следећим изразима:

$$Z_{RNAP}^{CR} = [RNAP \sim DNA] / [DNA] \quad (3.12)$$

$$Z_D^{CR} = [D \sim DNA] / [DNA] \quad (3.13)$$

$$Z_{D \sim RNAP}^{CR} = [D \sim DNA \sim RNAP] / [DNA] \quad (3.14)$$

$$Z_T^{CR} = [T \sim DNA] / [DNA] \quad (3.15)$$

док статистичка тежина празног промоторског региона износи 1. Увођењем израза за статистичке тежине (J-не (3.12)-(3.15)) у једначину (3.11) и коришћењем израза за равнотежне константе дисоцијације (J-не (3.6)-(3.10)) добија се коначни израз за зависност транскрипционе активности P.CR од концентрације мономера С протеина,

$$\varphi_{cr}(\text{Mon}) = \alpha \frac{a + b[\text{Mon}]^2}{1 + a + f[\text{Mon}]^2 + b[\text{Mon}]^2 + c[\text{Mon}]^4}, \quad (3.16)$$

са мањим бројем параметара који апсорбују константе (међу којима је и концентрација РНК полимеразе у ћелији која се сматра константном будући да је довољно велика да се не мења значајно услед везивања за P.CR):

$$a = [RNAP] / K_2 = k^{-1} [RNAP] e^{-\Delta G_{P.CR}} \quad (3.17)$$

$$f = 1 / (K_1 K_3) = k^{-2} e^{-\Delta G_D - \Delta G_{DBS}} \quad (3.18)$$

$$b = [RNAP] / K_1 K_3 K_5 = k^{-3} [RNAP] \omega' e^{-\Delta G_D - \Delta G_{DBS} - \Delta G_{P.CR}} \quad (3.19)$$

$$c = 1 / (K_1^2 K_3 K_4) = k^{-4} \omega e^{-2\Delta G_D - \Delta G_{DBS} - \Delta G_{PBS}} \quad (3.20)$$

Чланови који мере јачину кооперативности у везивању два С димера за ДНК (J-на (3.20)), односно јачину интеракције С димера и РНК полимеразе приликом њеног регрутовања на промотор (J-на (3.19)), су издвојени у посебне параметре $\omega = \exp(-\Delta G_T)$ и $\omega' = \exp(-\Delta G_{D \sim RNAP})$ (Rodic, Blagojevic, Zdobnov, et al., 2017).

Како метилтрансфераза својим деловањем утишава транскрипцију са промотора Р.М, претпостављено је да овај промотор може бити (Слика 3В) (Bogdanova et al., 2008; Rodic, Blagojevic, Zdobnov, et al., 2017):

- 1) празан у концентрацији [DNA],
- 2) активан у концентрацији [RNAP~DNA], у режиму базалне транскрипције када је за њега везана РНК полимеразе, или
- 3) утишан у концентрацији [M~DNA], када се за ДНК веже метилтрансфераза (концентрације [M]) која метилује секвенцу у оквиру промотора.

Према томе, регулацију Р.М описују следеће равнотежне реакције везивања:



чије су равнотежне константе дисцијације дате изразима:

$$K_{1M} = ([RNAP][DNA])/[RNAP\sim DNA] = k e^{\Delta G_{P,M}} \quad (3.23)$$

$$K_{2M} = ([M][DNA])/[M\sim DNA] = k e^{\Delta G_{Met}} \quad (3.24)$$

где су $\Delta G_{P,M}$ и ΔG_{Met} слободне енергије интеракције РНК полимеразе и метилтрансферазе са ДНК. Помоћу Ј-на (3.23) и (3.24) се могу изразити статистичке тежине конфигурација промотора Р.М:

$$Z_{RNAP}^M = [RNAP\sim DNA]/[DNA] \quad (3.25)$$

$$Z_M^M = [M\sim DNA]/[DNA] \quad (3.26)$$

на основу којих се одређује вероватноћа заузетости промотора РНК полимеразом у једначини за транскрипциону активност Р.М,

$$\varphi_m = \beta \frac{Z_{RNAP}^M}{1 + Z_{RNAP}^M + Z_M^M}, \quad (3.27)$$

где је β константа пропорционалности са јединицама [концентрација]/[време]. Коначни облик израза за транскрипциону активност промотора Р.М,

$$\varphi_m(M) = \varphi_m^{basal} \frac{K_D}{K_D + [M]}, \quad (3.28)$$

се добија увођењем следећа два параметра:

$$\varphi_m^{basal} = \beta \frac{[RNAP]}{K_{1M} + [RNAP]} \quad (3.29)$$

$$K_D = \left(1 + \frac{[RNAP]}{K_{1M}}\right) K_{2M} \quad (3.30)$$

где φ_m^{basal} представља његову базалну транскрипциону активност.

3.1.2 Динамичко моделовање експресије транскрипата и протеина система AhdI

Промена концентрација информационих РНК и протеина система са временом је описана системом обичних диференцијалних једначина (Bogdanova et al., 2008; Rodic, Blagojevic, Zdobnov, et al., 2017). Синтеза транскрипата *cr* оперона и *m* гена је дата изразима (3.16) и (3.28) термодинамичког модела транскрипције, док њихове концентрације опадају услед деградације, са константама деградације λ_{cr} и λ_m :

$$\frac{dcr(t)}{dt} = \varphi_{cr}(Mon(t)) - \lambda_{cr} \cdot cr(t) \quad (3.31)$$

$$\frac{dm(t)}{dt} = \varphi_m(M(t)) - \lambda_m \cdot m(t). \quad (3.32)$$

С протеин, рестрикциона ендонуклеаза и метилтрансфераза се у моделу синтетичу транслацијом одговарајућих транскрипата са константама транслације k_C , k_R и k_M , а елиминишу деградацијом са константама λ_C , λ_R и λ_M :

$$\frac{dC(t)}{dt} = k_C \cdot cr(t) - \lambda_C \cdot C(t) \quad (3.33)$$

$$\frac{dR(t)}{dt} = k_R \cdot cr(t) - \lambda_R \cdot R(t) \quad (3.34)$$

$$\frac{dM(t)}{dt} = k_M \cdot m(t) - \lambda_M \cdot M(t) \quad (3.35)$$

Будући да се и С протеин и рестрикциона ендонуклеаза синтетишу транслацијом cr транскрипта, уз претпоставку о једнаким константама деградације за ове протеине диференцијална једначина (3.34) се може заменити алгебарском:

$$R(t) = \frac{k_R}{k_C} C(t). \quad (3.36)$$

3.1.3 Одређивање вредности параметара модела Р-М система AhdI

Транскрипциона активност промотора Р.СR у *wild type* систему AhdI је експериментално измерена *in vitro* у присуству различитих концентрација С протеина (Bogdanova et al., 2008). То је омогућило да се одреде вредности параметара a , b и c фитовањем једначине оригиналног термодинамичког модела регулације Р.СR за $\alpha = 1$, која се од Ј-не (3.16) разликовала по томе што није садржала члан $f[\text{Mon}]^2$. Будући да се *wt* AhdI одликује високом равнотежном константом дисоцијације за реакцију димеризације С протеина, узето је да је $[C] \approx [\text{Mon}]$, тј. да се сви С протеини у раствору налазе у форми мономера (Rodic, Blagojevic, Zdobnov, et al., 2017). Вредност параметра

$$f = 1/(K_1 K_3) = \sqrt{c(K_4/K_3)} \quad (3.37)$$

је процењена на основу измерених вредности константи дисоцијације (McGeehan et al., 2006). Применом услова да транскрипционе активности $\varphi_{cr}(\text{Mon}(t))$ (Ј-на (3.16)) и $\varphi_m(M(t))$ (Ј-на (3.28)) у равнотежном стању буду једнаке и износе 1 nM/min (Bogdanova et al., 2008), подешене су вредности параметара α , K_D и φ_m^{basal} . Осталим параметрима динамичког модела су додељене генеричке, стандардне вредности (Phillips et al., 2012): полуживот транскрипта износи 5 минута, а полуживот протеина 30 минута. С обзиром на то да транскрипт гена за С протеин не поседује место за везивање рибозома, што узрокује његову ређу транслацију у односу на транскрипте осталих гена, задато је $k_M = k_R = 5k_C$. Вредности свих параметара су представљене у Табели 1.

Табела 1. Вредности параметара модела Р-М система AhdI.

Константа	Опис	Вредност	
α	Константа пропорционалности за Р.СR	1.7	nM/min
φ_m^{basal}	Базална транскрипциона активност Р.М	1.7	nM/min
a		$1.6 \cdot 10^{-1}$	
f	Константе које апсорбују релевантне слободне енергије протеин-протеин и протеин-ДНК	$2.8 \cdot 10^{-4}$	nM ⁻²
b	интеракција и концентрацију слободне РНК	$1.5 \cdot 10^{-1}$	nM ⁻²
c	полимеразе у ћелији	$1.2 \cdot 10^{-5}$	nM ⁻⁴
K_D		$6.5 \cdot 10^2$	nM
k_C	Константа транслације за С протеин	$6.0 \cdot 10^{-1}$	1/min
k_R	Константа транслације за рестрикциону ендонуклеазу	3.0	1/min

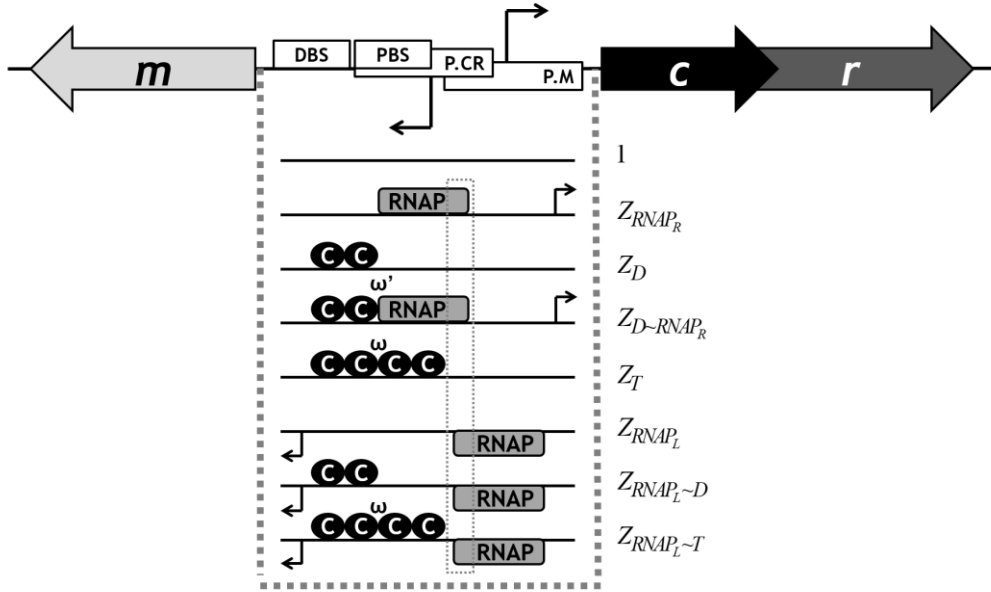
k_M	Константа транслације за метилтрансферазу	3.0	1/min
λ_{cr}	Константа деградације <i>cr</i> транскрипта	$2.0 \cdot 10^{-1}$	1/min
λ_m	Константа деградације <i>m</i> транскрипта	$2.0 \cdot 10^{-1}$	1/min
λ_C	Константа деградације С протеина	$3.3 \cdot 10^{-2}$	1/min
λ_R	Константа деградације рестрикционе ендонуклеазе	$3.3 \cdot 10^{-2}$	1/min
λ_M	Константа деградације метилтрансферазе	$3.3 \cdot 10^{-2}$	1/min

3.2 Биофизичко моделовање Р-М система EсoRV

3.2.1 Термодинамичко моделовање регулације транскрипције гена система EсoRV

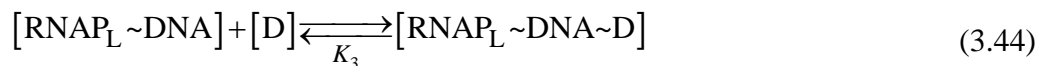
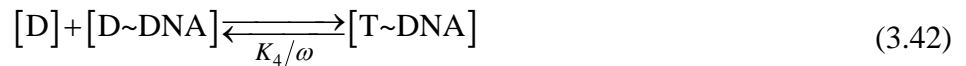
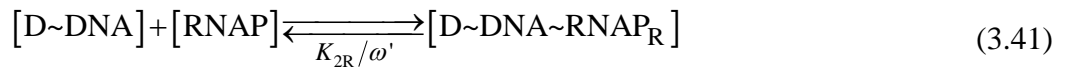
Будући да се Р-М систем EсoRV одликује дивергентном архитектуром, односно делимично преклопљеним промоторима Р.СR и Р.М за које РНК полимеразе не могу бити истовремено везане, први корак у моделовању је дефинисање дозвољених конфигурација комплетног интергенског региона (Слика 4). Претпостављено је да се интергенски регион може наћи у следећим конфигурацијама у назначеним концентрацијама (Rodic, Blagojevic, Zdobnov, et al., 2017):

- 1) [DNA] – конфигурација празне интергенске ДНК;
- 2) [RNAP_R~DNA] – конфигурација која настаје самосталним везивањем РНК полимеразе за Р.СR чиме се омогућава базална транскрипција *cr* оперона (на слици усмерена надесно – енг. *Right*);
- 3) [T~DNA] – конфигурација у којој су два димера С протеина везана за ДНК и утишавају транскрипцију *cr* оперона тиме што спречавају везивање РНК полимеразе за Р.СR;
- 4) [D~DNA~RNAP_R] – конфигурација у којој је РНК полимеразе регрутована на Р.СR од стране С димера везаног за DBS који делује као активатор транскрипције *cr* оперона;
- 5) [D~DNA] – конфигурација у којој је само један С димер везан за ДНК, која је оправдана експериментално приказаним одсуством кооперативности у везивању С протеина у систему EсoRV;
- 6) [RNAP_L~DNA] – конфигурација у којој је РНК полимеразе самостално везана за Р.М чиме је омогућена базална транскрипција *m* гена (на слици усмерена налево – енг. *Left*);
- 7) [RNAP_L~DNA~D] – конфигурација у којој је осим РНК полимеразе на Р.М присутан и С димер везан за DBS;
- 8) [RNAP_L~DNA~T] – конфигурација у којој је РНК полимеразе везана за Р.М, уз истовремено присуство два С димера на својим везивним местима.



Слика 4. Графички приказ регулације транскрипције у P-M систему EcoRV. Организација гена система на плазмиду је приказана одговарајуће именованим широким стрелицама, док правоугаоници означени са P.CR P.M, DBS и PBS представљају, редом, промоторе за транскрипцију *cr* оперона и *m* гена и места за везивање C димера (дистално и проксимално). Танке стрелице указују на смер транскрипције са датих промотора. Приказане су дозвољене конфигурације РНК полимеразе (правоугаоник RNAP) и C протеина (C мономер представљен кругом) на ДНК интергенског региона. Конфигурације које садрже стрелицу су транскрипционо активне у назначеном смеру. Објашњења ознака ω , ω' и статистичких тежина наведених уз сваку конфигурацију су дата у главном тексту. Слика адаптирана из (Rodic et al., 2018).

Ове конфигурације настају у следећим равнотежним реакцијама везивања:



за које се могу написати равнотежне константе дисоцијације:

$$K_1 = [\text{Mon}]^2 / [\text{D}] \quad (3.46)$$

$$K_{2R} = ([\text{RNAP}] [\text{DNA}]) / [\text{RNAP}_R \sim \text{DNA}] \quad (3.47)$$

$$K_3 = ([\text{D}] [\text{DNA}]) / [\text{D} \sim \text{DNA}] = ([\text{RNAP}_L \sim \text{DNA}] [\text{D}]) / [\text{RNAP}_L \sim \text{DNA} \sim \text{D}] \quad (3.48)$$

$$K_{2R} / \omega' = ([\text{D} \sim \text{DNA}] [\text{RNAP}]) / [\text{D} \sim \text{DNA} \sim \text{RNAP}_R] \quad (3.49)$$

$$K_4 / \omega = ([\text{D}] [\text{D} \sim \text{DNA}]) / [\text{T} \sim \text{DNA}] = ([\text{RNAP}_L \sim \text{DNA} \sim \text{D}] [\text{D}]) / [\text{RNAP}_L \sim \text{DNA} \sim \text{T}] \quad (3.50)$$

$$K_{2L} = ([\text{RNAP}] [\text{DNA}]) / [\text{RNAP}_L \sim \text{DNA}]. \quad (3.51)$$

Модел претпоставља да је везивање С протеина за ДНК независно од везивања РНК полимеразе за Р.М (видети реакције (3.40) и (3.44) и Ј-ну (3.48), као и реакције (3.42) и (3.45) и Ј-ну (3.50)) (Rodic, Blagojevic, Zdobnov, et al., 2017). Као и приликом моделовања система AhdI (поглавље 3.1.1), уведени су параметри $\omega = \exp(-\Delta G_T)$ за ниво кооперативности у везивању димера С протеина и $\omega' = \exp(-\Delta G_{D \sim \text{RNAP}})$ за јачину интеракције С димера везаног за DBS са РНК полимеразом регрутованом на Р.СР. Ради смањивања броја параметара модела, равнотежна константа дисоцијације реакције у којој се РНК полимеразе регрутује на Р.СР (Ј-на (3.49)) је изражена као однос константе за реакцију везивања РНК полимеразе за Р.СР (K_{2R}) и ω' . Слично, равнотежна константа дисоцијације за реакцију формирања С тетрамера на ДНК (Ј-на (3.50)) је изражена као однос константе за везивање С димера за PBS (K_4) и ω .

Транскрипционе активности Р.СР и Р.М,

$$\varphi_{cr} = \gamma \frac{Z_{\text{RNAP}_R} + Z_{\text{D} \sim \text{RNAP}_R}}{1 + Z_{\text{RNAP}_R} + Z_D + Z_T + Z_{\text{D} \sim \text{RNAP}_R} + Z_{\text{RNAP}_L} + Z_{\text{RNAP}_L \sim \text{D}} + Z_{\text{RNAP}_L \sim \text{T}}} \quad (3.52)$$

$$\varphi_m = \gamma \frac{Z_{\text{RNAP}_L} + Z_{\text{RNAP}_L \sim \text{D}} + Z_{\text{RNAP}_L \sim \text{T}}}{1 + Z_{\text{RNAP}_R} + Z_D + Z_T + Z_{\text{D} \sim \text{RNAP}_R} + Z_{\text{RNAP}_L} + Z_{\text{RNAP}_L \sim \text{D}} + Z_{\text{RNAP}_L \sim \text{T}}} \quad (3.53)$$

су пропорционалне равнотежним вероватноћама заузетости ових промотора РНК полимеразом (Shea and Ackers, 1985), где је γ константа пропорционалности са јединицама [концентрација]/[време]. Ове вероватноће су изведене преко статистичких тежина конфигурација интергенског региона (Слика 4):

$$Z_{\text{RNAP}_R} = [\text{RNAP}_R \sim \text{DNA}] / [\text{DNA}] \quad (3.54)$$

$$Z_D = [\text{D} \sim \text{DNA}] / [\text{DNA}] \quad (3.55)$$

$$Z_T = [\text{T} \sim \text{DNA}] / [\text{DNA}] \quad (3.56)$$

$$Z_{\text{D} \sim \text{RNAP}_R} = [\text{D} \sim \text{DNA} \sim \text{RNAP}_R] / [\text{DNA}] \quad (3.57)$$

$$Z_{\text{RNAP}_L} = [\text{RNAP}_L \sim \text{DNA}] / [\text{DNA}] \quad (3.58)$$

$$Z_{\text{RNAP}_L \sim \text{D}} = [\text{RNAP}_L \sim \text{DNA} \sim \text{D}] / [\text{DNA}] \quad (3.59)$$

$$Z_{\text{RNAP}_L \sim \text{T}} = [\text{RNAP}_L \sim \text{DNA} \sim \text{T}] / [\text{DNA}]. \quad (3.60)$$

Коначне једначине за транскрипционе активности промотора:

$$\varphi_{cr}(\text{Mon}) = \gamma \frac{\left(1 + \omega \frac{[\text{Mon}]^2}{K_1 K_3}\right)}{x \left(1 + \frac{[\text{Mon}]^2}{K_1 K_3} + \omega \frac{[\text{Mon}]^4}{5 K_1^2 K_3^2}\right) + \left(1 + \omega \frac{[\text{Mon}]^2}{K_1 K_3}\right)} \quad (3.61)$$

$$\varphi_m(\text{Mon}) = \gamma \frac{x \left(1 + \frac{[\text{Mon}]^2}{K_1 K_3} + \omega \frac{[\text{Mon}]^4}{5 K_1^2 K_3^2}\right)}{x \left(1 + \frac{[\text{Mon}]^2}{K_1 K_3} + \omega \frac{[\text{Mon}]^4}{5 K_1^2 K_3^2}\right) + \left(1 + \omega \frac{[\text{Mon}]^2}{K_1 K_3}\right)} \quad (3.62)$$

су добијене увођењем израза (3.54)-(3.60) и (3.46)-(3.51) у Ј-не (3.52) и (3.53). Притом је узето у обзир да DBS има пет пута већи афинитет за везивање С протеина него PBS, тј. $K_4/K_3 = 5$, као и да је P.CR знатно слабији промотор од P.M, тј. $K_{2R} \gg K_{2L}$ (Semenova et al., 2005). Такође, уведен је нови параметар

$$x = \frac{[\text{RNAP}]}{K_{2L}} \cdot \left(\frac{[\text{RNAP}]}{K_{2R}}\right)^{-1}. \quad (3.63)$$

3.2.2 Динамичко моделовање експресије транскрипата и протеина система EcoRV

Термодинамички модел регулације транскрипције (Ј-не (3.61) и (3.62)) је уклопљен у динамички модел експресије cr и m транскрипата, С протеина, рестрикционе ендонуклеазе и метилтрансферазе (Rodic, Blagojevic, Zdobnov, et al., 2017). Као и у случају система AhdI, диференцијална једначина за промену концентрације рестрикционе ендонуклеазе са временом је замењена алгебарском,

$$R(t) = \frac{k_R}{k_C} C(t), \quad (3.64)$$

што важи под претпоставком да су константе деградације С протеина и рестрикционе ендонуклеазе једнаке, па се динамички модел састоји из четири обичне диференцијалне једначине:

$$\frac{dcr(t)}{dt} = \varphi_{cr}(\text{Mon}(t)) - \lambda_{cr} \cdot cr(t) \quad (3.65)$$

$$\frac{dm(t)}{dt} = \varphi_m(\text{Mon}(t)) - \lambda_m \cdot m(t) \quad (3.66)$$

$$\frac{dC(t)}{dt} = k_C \cdot cr(t) - \lambda_C \cdot C(t) \quad (3.67)$$

$$\frac{dM(t)}{dt} = k_M \cdot m(t) - \lambda_M \cdot M(t). \quad (3.68)$$

Ј-не (3.65) и (3.66) описују промену концентрација транскрипата са временом, који се синтетишу транскрипцијом одговарајућих гена, а елиминишу деградацијом. Протеини се синтетишу транслацијом транскрипата, а уклањају из система деградацијом, што описују Ј-не (3.67) и (3.68).

3.2.3 Рескалирање величина и одређивање вредности параметара модела Р-М система EcoRV

Wild type EcoRV карактеришу релативно ниска равнотежна константа дисоцијације за реакцију димеризације С протеина и одсуство кооперативности у везивању С димера за DBS и PBS (Semenova et al., 2005). У складу са првим, транскрипционе активности Р.СR (Ј-на (3.61)) и Р.М (Ј-на (3.62)) су изражене у функцији концентрације С димера коришћењем једначине (3.46):

$$\varphi_{cr}(D) = \gamma \frac{\left(1 + \omega' \frac{[D]}{K_3}\right)}{x \left(1 + \frac{[D]}{K_3} + \omega \frac{[D]^2}{5K_3^2}\right) + \left(1 + \omega' \frac{[D]}{K_3}\right)} \quad (3.69)$$

$$\varphi_m(D) = \gamma \frac{x \left(1 + \frac{[D]}{K_3} + \omega \frac{[D]^2}{5K_3^2}\right)}{x \left(1 + \frac{[D]}{K_3} + \omega \frac{[D]^2}{5K_3^2}\right) + \left(1 + \omega' \frac{[D]}{K_3}\right)}. \quad (3.70)$$

На основу претпоставке да су сви протеини $[C] = [Mon] + 2[D]$ у раствору у форми димера постављен је услов $[D] \approx [C]/2$. Коришћењем тог услова, дељењем обе стране једначина (3.69) и (3.70) са $2K_3$ и увођењем рескалираних величина $\bar{C} = [C]/(2K_3)$, $\bar{\gamma} = \gamma/(2K_3)$, $\bar{\varphi}_{cr} = \varphi_{cr}/(2K_3)$ и $\bar{\varphi}_m = \varphi_m/(2K_3)$ добијене су следеће једначине:

$$\bar{\varphi}_{cr}(\bar{C}) = \bar{\gamma} \frac{(1 + \omega' \bar{C})}{x(1 + \bar{C} + \omega \frac{\bar{C}^2}{5}) + (1 + \omega' \bar{C})} \quad (3.71)$$

$$\bar{\varphi}_m(\bar{C}) = \bar{\gamma} \frac{x(1 + \bar{C} + \omega \frac{\bar{C}^2}{5})}{x(1 + \bar{C} + \omega \frac{\bar{C}^2}{5}) + (1 + \omega' \bar{C})}. \quad (3.72)$$

Ради усклађивања динамичког модела, Ј-не (3.65)-(3.68) и (3.64) су такође подељене са $2K_3$ и концентрације замењене рескалираним $\bar{c}_r = [cr]/(2K_3)$, $\bar{m} = [m]/(2K_3)$, $\bar{M} = [M]/(2K_3)$ и $\bar{R} = [R]/(2K_3)$:

$$\frac{d\bar{c}_r(t)}{dt} = \bar{\varphi}_{cr}(\bar{C}(t)) - \lambda_{cr} \cdot \bar{c}_r(t) \quad (3.73)$$

$$\frac{d\bar{m}(t)}{dt} = \bar{\varphi}_m(\bar{C}(t)) - \lambda_m \cdot \bar{m}(t) \quad (3.74)$$

$$\frac{d\bar{C}(t)}{dt} = k_C \cdot \bar{c}_r(t) - \lambda_C \cdot \bar{C}(t) \quad (3.75)$$

$$\frac{d\bar{M}(t)}{dt} = k_M \cdot \bar{m}(t) - \lambda_M \cdot \bar{M}(t) \quad (3.76)$$

$$\bar{R}(t) = \frac{k_R}{k_C} \bar{C}(t). \quad (3.77)$$

Рескалирањем величина је смањен број параметара модела чије се вредности морају проценити, при чему модел и даље може да се користи за испитивање зависности особина динамике експресије система од његових конструктивних карактеристика, будући да апсолутне вредности концентрација молекула нису релевантне за ову анализу (Rodic, Vladojevic, Zdobnov, et al., 2017). У складу са одсуством кооперативности у везивању два С димера за ДНК, у *wt* систему је $\omega = 1$. Вредност параметра $x = 10$ је фиксирана под претпоставком о 10 пута већој базалној транскрипционој активности Р.М у односу на Р.СР (Semenova et al., 2005), док је фиксирањем $\omega' = 20$ одабрана за ред величине мања јачина интеракције између С димера и РНК полимеразе од измерене за други Р-М систем са кооперативним везивањем два С димера (Bogdanova et al., 2009). У равнотежном стању, први изводи концентрација по времену (Ј-не 3.73-3.76) су једнаки 0. Примена услова о једнакости равнотежних концентрација *cr* и *m* транскриптата, у складу са експерименталним запажањем из (Bogdanova et al., 2008), ако се задају исте вредности λ_{cr} и λ_m , води захтеву за једнаким транскрипционим активностима промотора (Ј-не (3.71) и (3.72)) у равнотежном стању, односно једнакости

$$1 + \omega' \bar{C}_{eq} = x(1 + \bar{C}_{eq} + \frac{\bar{C}_{eq}^2}{5}) \quad (3.78)$$

из које се може одредити равнотежна концентрација С протеина, \bar{C}_{eq} . Затим се може одредити транскрипциона активност Р.СР у равнотежном стању,

$$\bar{\varphi}_{cr}(\bar{C}_{eq}) = \frac{\lambda_{cr} \lambda_C}{k_C} \bar{C}_{eq} \quad (3.79)$$

применом услова за равнотежу на Ј-не (3.73) и (3.75) и додељивањем стандардних вредности параметрима (Phillips et al., 2013), слично као приликом моделовања система AhdI. Конкретно, фиксирани су вредности $\lambda_{cr} = \lambda_m = 1/5$ 1/min, $\lambda_C = \lambda_R = \lambda_M = 1/30$ 1/min, као и $k_C = k_R = k_M = 3$ 1/min пошто у систему EcoRV транскрипти свих гена имају место за везивање рибозома (Semenova et al., 2005). Коначно, вредност параметра $\bar{\gamma}$ се може одредити из Ј-не (3.71) за равнотежно стање.

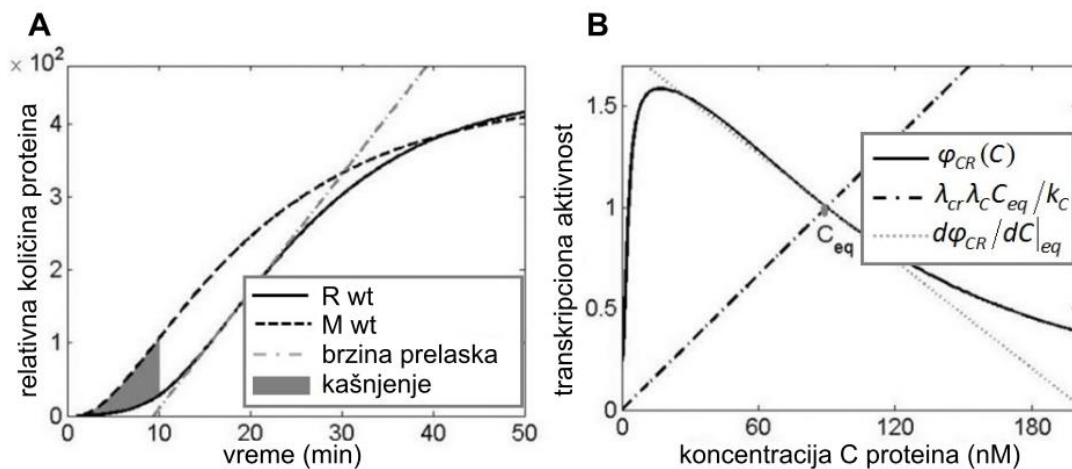
3.3 Мерење ефеката ремећења специфичних конструктивних карактеристика Р-М система AhdI и EcoRV на својства динамике експресије

3.3.1 Дефинисање квантитативних мера за степен испољености жељених својстава динамике експресије Р-М система

Ефекти ремећења специфичних конструктивних карактеристика Р-М система AhdI и EcoRV су описани праћењем промена три својства динамике експресије Р-М система која су потенцијално важна за обављање њихове функције, а то су (Rodic, Vladojevic, Zdobnov, et al., 2017):

- 1) кашњење почетка експресије рестрикционе ендонуклеазе којим се обезбеђује акумулација довољне количине метилтрансферазе,
- 2) брз прелазак експресије рестрикционе ендонуклеазе из искљученог у укључено стање када почне њена синтеза и
- 3) одсуство великог флукутирања концентрација протеина у равнотежном стању које би могло да доведе до неизбалансиране ендонуклеазне и метилтрансферазне активности у ћелији.

За свако од ових динамичких својстава је дефинисана квантитативна мера чија се вредност може одредити за различите пертурбације параметара модела. Дефиниције мера су графички објашњене на примеру динамике раније моделованог система AhdI (Слика 5) (Bogdanova et al., 2008).



Слика 5. Графички приказ величина које мере испољеност динамичких својстава P-M система на примеру експресије протеина система AhdI приликом његовог успостављања у новој ћелији-домаћину. (A) Динамика промене релативних нивоа рестрикционе ендонуклеазе (R) и метилтрансферазе (M) је приказана црним кривама од момента уласка плаزمида у ћелију који одговара нултом времену. Осенчена површина је узета као мера кашњења експресије R у односу на M, док је брзина преласка експресије R из искљученог у укључено стање дата нагибом светло сиве праве. (B) Испрекидана линија приказује нагиб криве зависности транскрипционе активности P.CR од концентрације C протеина у тачки која одговара равнотежној концентрацији. Овај нагиб се користи за израчунавање мере стабилности стационарног стања система. Слика адаптирана из (Rodic, Blagojevic, Zdobnov, et al., 2017).

Мера кашњења експресије рестрикционе ендонуклеазе. Кашњење експресије рестрикционе ендонуклеазе се мери површином између кривих за метилтрансферазу и рестрикциону ендонуклеазу на графику динамике експресије протеина, за првих 10 минута по уласку система на плазмиду у нову бактеријску ћелију (осенчена површина на Слици 5A). Приликом варирања одабраног параметра модела, одређује се однос ових површина у *in silico* мутираном и *wt* систему, тј. „релативно кашњење“:

$$\tau_{delay} = \left(\int_1^{10} M_{mut}(t)dt - \int_1^{10} R_{mut}(t)dt \right) / \left(\int_1^{10} M_{wt}(t)dt - \int_1^{10} R_{wt}(t)dt \right). \quad (3.80)$$

Према Ј-ни (3.80), за *wt* систем је $\tau_{delay} = 1$, тако да $\tau_{delay} < 1$ указује на то да дата пертурбација параметра доводи до скраћивања кашњења у експресији рестрикционе ендонуклеазе, и обратно.

Мера брзине преласка експресије рестрикционе ендонуклеазе из искљученог у укључено стање. Крива R(t) има сигмоидни облик: за мала времена концентрације протеина су ниске, за велика високе, док у средњем интервалу концентрација релативно брзо расте. Брзину преласка експресије рестрикционе ендонуклеазе из искљученог у укључено стање (v_{max}), стога, даје максимални нагиб R(t) криве у средњем временском интервалу (тачка-цртаста крива на Слици 5A).

Мера стабилности система у равнотежном стању. Ако се на диференцијалне једначине за промену концентрација *cr* транскрипта и C протеина,

$$\frac{dcr(t)}{dt} = \varphi_{cr}(C(t)) - \lambda_{cr} \cdot cr(t) \quad (3.81)$$

$$\frac{dC(t)}{dt} = k_C \cdot cr(t) - \lambda_C \cdot C(t), \quad (3.82)$$

какве чине динамичке моделе AhdI (J-не (3.31) и (3.33)) и EcoRV (J-не (3.65) и (3.67)), примени услов за равнотежу $dcr(t)/dt = dC(t)/dt = 0$, добија се линеарна зависност равнотежне транскрипционе активности P.CR од концентрације C протеина:

$$\varphi_{cr}(C_{eq}) = \frac{\lambda_{cr} \lambda_C}{k_C} C_{eq}. \quad (3.83)$$

На Слици 5B је ова зависност представљена правом са нагибом $\lambda_{cr} \lambda_C / k_C$. Да би $\varphi_{cr}(C_{eq})$ било стабилно стационарно стање, мале промене у концентрацији C протеина које је изводе из равнотежне вредности морају бити праћене враћањем система у то стање. Ради испитивања стабилности равнотежног стања, изведене су једначине за динамику система у његовој околини увођењем $cr(t) = cr_{eq} + \delta cr(t)$ и $C(t) = C_{eq} + \delta C(t)$ у J-не (3.81) и (3.82), где су $\delta cr(t)$ и $\delta C(t)$ мале вредности концентрација за које су промењене равнотежне концентрације cr транскрипта и C протеина. Како је $dcr_{eq}/dt = dC_{eq}/dt = 0$ и узимајући да је

$$\varphi_{cr}(C_{eq} + \delta C(t)) = \varphi_{cr}(C_{eq}) + \left. \frac{d\varphi_{cr}(C)}{dC} \right|_{C_{eq}} \delta C(t), \quad (3.84)$$

добијене су следеће једначине:

$$\frac{d\delta cr(t)}{dt} = \left. \frac{d\varphi_{cr}(C)}{dC} \right|_{C_{eq}} \delta C(t) - \lambda_{cr} \cdot \delta cr(t) \quad (3.85)$$

$$\frac{d\delta C(t)}{dt} = k_C \cdot \delta cr(t) - \lambda_C \cdot \delta C(t). \quad (3.86)$$

Даље, $\delta cr(t)$ може да се изрази из J-не (3.86) и да се убаци у J-ну (3.85), чиме се добија једначина пригушеног хармонијског осцилатора,

$$\frac{d^2 \delta C(t)}{dt^2} + (\lambda_{cr} + \lambda_C) \frac{d\delta C(t)}{dt} + \left(\lambda_{cr} \cdot \lambda_C - k_C \left. \frac{d\varphi_{cr}(C)}{dC} \right|_{C_{eq}} \right) \cdot \delta C(t) = 0, \quad (3.87)$$

чији коефицијент уз линеарни члан (пропорционалан константи/крутости опруге) одговара мери стабилности стационарног стања:

$$\Omega^2 = \lambda_{cr} \cdot \lambda_C - k_C \left. \frac{d\varphi_{cr}(C)}{dC} \right|_{C_{eq}}. \quad (3.88)$$

Позитивна вредност Ω^2 одликује стабилно стационарно стање система, а добија се када је нагиб праве зависности равнотежне транскрипционе активности P.CR од концентрације C протеина, $\lambda_{cr} \lambda_C / k_C$, већи од нагиба криве $\varphi_{cr}(C)$ у тачки њиховог пресека која одговара равнотежној концентрацији C протеина, C_{eq} (Слика 5B).

3.3.2 Пертурбисање равнотежне константе дисоцијације за реакцију димеризације С протеина

Пертурбисање K_1 у систему AhdI. Варирањем вредности равнотежне константе дисоцијације за реакцију димеризације (3.1), K_1 , у моделу система AhdI се мења однос концентрација мономера и димера С протеина у раствору, као и вредности константи f , b и c (J-не (3.18)-(3.20)). Како је укупна концентрација синтетисаног С протеина $[C] = [Mon] + 2[D]$, а $K_1 = [Mon]^2/[D]$ према J-ни (3.6), концентрација С мономера је одређена следећим изразом:

$$[Mon] = \frac{K_1}{4} \left(\sqrt{1 + \frac{8[C]}{K_1}} - 1 \right). \quad (3.89)$$

Гранични случајеви реакције димеризације С протеина подразумевају:

- 1) налажење свих С протеина у раствору у форми мономера, $[C] \approx [Mon]$, када је вредност K_1 веома висока, и
- 2) присуство само С димера у раствору, $[C] \approx 2[D]$, за веома ниску вредност K_1 .

Први случај приближно одликује *wt* систем AhdI чија је динамика експресије симулирана нумеричким решавањем J-на (3.16), (3.28), (3.31)-(3.33) и (3.35)-(3.36) уз претпоставку да је $[C] \approx [Mon]$ и вредности параметара наведене у поглављу 3.1.3. Затим је уместо апроксимације $[C] \approx [Mon]$ коришћена J-на (3.89) и динамика је симулирана за различите вредности K_1 , док није идентификована најмања вредност при којој се може добити крива $R(t)$ која не одступа значајно од исте добијене симулацијом динамике *wt* система. Ова вредност ($K_1^{max} \approx 4000$ nM) је узета као вредност константе дисоцијације за реакцију димеризације у *wt* систему AhdI (Rodic, Vladojevic, Zdobnov, et al., 2017).

У другом граничном случају, за $K_1 \ll 8[C]$, J-на (3.89) се своди на $[Mon] \approx \sqrt{K_1[C]/2}$, што применом у J-ни (3.6) води управо услову $[C] \approx 2[D]$. Ако се *wt* вредност K_1 смањи n пута, тако да је $K_1 = K_1^{max}/n \ll 8[C]$, и конзистентно се измене вредности параметара $f = f_{wt} \cdot n$, $b = b_{wt} \cdot n$ и $c = c_{wt} \cdot n^2$ у односу на вредности фиксиране за *wt* систем (Табела 1), транскрипциона активност P.CR (J-на (3.16)) у овом граничном случају постаје:

$$\varphi_{cr}(C) = \alpha \frac{a + \frac{b}{2} K_1^{max} [C]}{1 + a + \frac{f}{2} K_1^{max} [C] + \frac{b}{2} K_1^{max} [C] + \frac{c}{4} (K_1^{max})^2 [C]^2}. \quad (3.90)$$

Односи концентрација мономера и димера између ових граничних случајева се добијају варирањем n од 1 до максималне вредности за коју се крива зависности $\varphi_{cr}(C)$ преклапа са кривом коју одређује J-на (3.90), при чему се J-на (3.89) користи за добијање концентрације С мономера.

Пертурбисање K_1 у систему EcoRV. У рескалираној верзији модела EcoRV, слично као у оригиналној, важи $\bar{C} = \overline{Mon} + 2\bar{D}$, где су $\overline{Mon} = [Mon]/(2K_3)$ и $\bar{D} = [D]/(2K_3)$ рескалирани као и друге величине у поглављу 3.2.3. Из ове и рескалиране једначине (3.46), $\overline{\overline{K_1}} = \overline{Mon}^2/(8\bar{D})$, са нешто другачије дефинисаним $\overline{\overline{K_1}} = K_1/(16K_3)$, добија се:

$$\overline{Mon} = 2\overline{\overline{K_1}} \left(\sqrt{1 + \frac{\bar{C}}{\overline{\overline{K_1}}}} - 1 \right). \quad (3.91)$$

Транскрипциону активност промотора у *wt* систему описују J-не (3.71) и (3.72) које апроксимирају гранични случај у коме се сви С протеини у раствору налазе у форми димера, $[C] \approx 2[D]$, док се општији облик добија рескалирањем J-на (3.61)-(3.62):

$$\bar{\varphi}_{cr}(\overline{\text{Mon}}) = \bar{\gamma} \frac{\left(1 + \omega \frac{\overline{\text{Mon}}^2}{4K_1}\right)}{x \left(1 + \frac{\overline{\text{Mon}}^2}{4K_1} + \omega \frac{\overline{\text{Mon}}^4}{80K_1}\right) + \left(1 + \omega \frac{\overline{\text{Mon}}^2}{4K_1}\right)} \quad (3.92)$$

$$\bar{\varphi}_m(\overline{\text{Mon}}) = \bar{\gamma} \frac{x \left(1 + \frac{\overline{\text{Mon}}^2}{4K_1} + \omega \frac{\overline{\text{Mon}}^4}{80K_1}\right)}{x \left(1 + \frac{\overline{\text{Mon}}^2}{4K_1} + \omega \frac{\overline{\text{Mon}}^4}{80K_1}\right) + \left(1 + \omega \frac{\overline{\text{Mon}}^2}{4K_1}\right)} \quad (3.93)$$

Вредност $\overline{K_1}^{min} \approx 1/1000$ је одређена подешавањем $\overline{K_1}$ у J-нама (3.91)-(3.93) тако да се криве R(t) и M(t) добијене решавањем динамичког модела (3.73)-(3.77) са овим једначинама поклопе са кривама *wt* система, добијеним решавањем динамичког модела са J-нама (3.71) и (3.72). Горња гранична вредност $\overline{K_1}^{max} \approx 10$ је одређена на сличан начин, тражењем кривих R(t) и M(t) које се преклапају са кривама модела који подразумева да су сви C протеини у раствору у форми мономера, а укључује J-не (3.92)-(3.93) уз услов $\overline{\text{Mon}} \approx \bar{C}$. Између ових граничних вредности је варирано $\overline{K_1}$ и притом је симулирана динамика система решавањем J-на (3.73)-(3.77) и (3.91)-(3.93) (Rodic, Blagojevic, Zdobnov, et al., 2017).

3.3.3 Пертурбисање нивоа кооперативности у везивању два C димера за ДНК и константе транслације за C протеин

Пертурбисање ω . Вредност јачине кооперативности у везивању два C димера за ДНК је процењена на $\omega \approx 3000$ за систем AhdI (McGeehan et al., 2006), па је претпостављено да *wt* вредност параметра *c* (J-на (3.20), Табела 1) подразумева ову вредност ω . Динамика система AhdI је симулирана и анализирана решавањем једначина модела (3.16), (3.28), (3.31)-(3.33) и (3.35)-(3.36) при различитим вредностима *c*, у опсегу ω од 3000 до 1, што одговара одсуству кооперативности.

У рескалираним једначинама модела EcoRV (J-не (3.71), (3.72) и (3.73)-(3.77)) је ω постепено повећавано почевши од 1, што одговара одсуству кооперативности у *wt* систему.

Пертурбисање k_C . Вредност константе транслације за C протеин у моделима система AhdI и EcoRV је варирана између 3/5 1/min, која одговара *wt* систему AhdI, и 3 1/min, као у *wt* систему EcoRV.

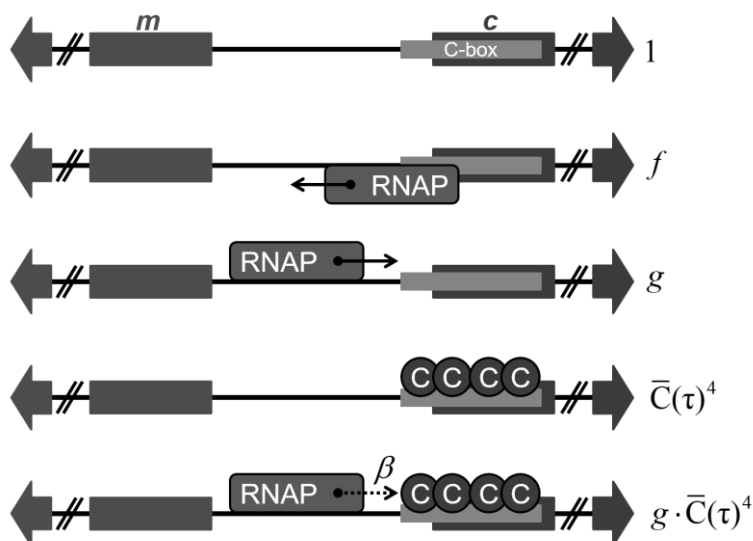
3.4 Биофизичко моделовање Р-М система Kрп2I

3.4.1 Термодинамичко моделовање регулације транскрипције гена система Kрп2I

У термодинамички модел регулације преклопљених промотора Р.М и Р.С су ушле следеће конфигурације (у назначеним концентрацијама) интергенског региона у коме се они налазе, продуженог тако да обухвата и део *c* гена са местима за везивање C димера (Слика 6) (Klimuk et al., 2018):

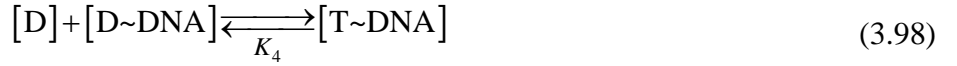
- 1) [DNA] – конфигурација празне ДНК;

- 2) $[RNAP_L \sim DNA]$ – конфигурација која омогућава базалну транскрипцију m гена (на слици усмерену налево – енг. *Left*), када је РНК полимераза самостално везана за Р.М и спречава истовремено везивање РНК полимеразе за Р.С;
- 3) $[RNAP_R \sim DNA]$ – конфигурација која настаје самосталним везивањем РНК полимеразе за Р.С, када се одвија базална транскрипција c гена (на слици усмерена надесно – енг. *Right*), а РНК полимераза притом не може да се веже за Р.М;
- 4) $[T \sim DNA]$ – конфигурација у којој су два димера С протеина везана за ДНК и спречавају истовремено везивање РНК полимеразе за Р.М утишавајући транскрипцију m гена (конфигурација у којој је само један С димер на ДНК није укључена јер је експериментално описана само улога тетрамера);
- 5) $[RNAP_R \sim DNA \sim T]$ – конфигурација у којој је РНК полимераза везана за Р.С док су два С димера на својим везивним местима (у овом случају долази до иницијације, али везани С тетрамер смањује ефикасност елонгације транскрипције).



Слика 6. Графички приказ регулације транскрипције у Р-М систему Крп2I. Приказане су дозвољене конфигурације РНК полимеразе (правоугаоник RNAP) и С протеина (С мономер представљен кругом) на ДНК у региону између супротно усмерених m и c гена представљених широким стрелицама. Смер транскрипције РНК полимеразе везане за промотор одређеног гена је назначен стрелицом, при чему испрекидана стрелица у конфигурацији на дну указује на мању ефикасност елонгације транскрипције коју мери параметар β . Претпостављено је да тетрамер С протеина може бити везан за С-box представљен правоугаоником који је делимично преклопљен са секвенцом c гена. Са десне стране су назначене статистичке тежине конфигурација чији су коначни изрази изведени у поглављу 3.4.3. Слика адаптирана из (Klimuk et al., 2018).

Описане конфигурације настају кроз следеће равнотежне реакције везивања са одговарајућим равнотежним константама дисоцијације (K):



и одговарају им следеће статистичке тежине:

$$Z_{\text{RNAP}_L} = \frac{[\text{RNAP}_L \sim \text{DNA}]}{[\text{DNA}]} = \frac{[\text{RNAP}]}{K_{2L}} = f \quad (3.100)$$

$$Z_{\text{RNAP}_R} = \frac{[\text{RNAP}_R \sim \text{DNA}]}{[\text{DNA}]} = \frac{[\text{RNAP}]}{K_{2R}} = g \quad (3.101)$$

$$Z_T = \frac{[\text{T} \sim \text{DNA}]}{[\text{DNA}]} = \frac{1}{K_1^2 K_3 K_4} [\text{Mon}]^4 = \frac{1}{h} [\text{Mon}]^4 \quad (3.102)$$

$$Z_{\text{RNAP}_R \sim \text{T}} = \frac{[\text{RNAP}_R \sim \text{DNA} \sim \text{T}]}{[\text{DNA}]} = \frac{[\text{RNAP}]}{K_{2R}} \cdot \frac{1}{K_1^2 K_3 K_4} [\text{Mon}]^4 = \frac{g}{h} [\text{Mon}]^4, \quad (3.103)$$

где су f , g и h параметри који замењују константне чланове у изразима за статистичке тежине. Према стандардној претпоставци о сразмерности стопе иницијације транскрипције вероватноћи налажења РНК полимеразе на промотору (Shea & Ackers, 1985), транскрипциона активност Р.М је дата изразом

$$\varphi_m = \alpha \frac{Z_{\text{RNAP}_L}}{1 + Z_{\text{RNAP}_L} + Z_{\text{RNAP}_R} + Z_T + Z_{\text{RNAP}_R \sim \text{T}}}, \quad (3.104)$$

док транскрипциона активност Р.С,

$$\varphi_c = \alpha \frac{Z_{\text{RNAP}_R} + Z_{\text{RNAP}_R \sim \text{T}}}{1 + Z_{\text{RNAP}_L} + Z_{\text{RNAP}_R} + Z_T + Z_{\text{RNAP}_R \sim \text{T}}} \cdot \left(\frac{1 + Z_{\text{RNAP}_L} + Z_{\text{RNAP}_R}}{1 + Z_{\text{RNAP}_L} + Z_{\text{RNAP}_R} + Z_T + Z_{\text{RNAP}_R \sim \text{T}}} + \beta \frac{Z_T + Z_{\text{RNAP}_R \sim \text{T}}}{1 + Z_{\text{RNAP}_L} + Z_{\text{RNAP}_R} + Z_T + Z_{\text{RNAP}_R \sim \text{T}}} \right) \quad (3.105)$$

представља стопу иницијације транскрипције (први ред Ј-не (3.105)) помножену вероватноћом да транскрипт гена буде до краја синтетисан (други ред Ј-не (3.105)). Ова вероватноћа је дата збиром два члана: 1) вероватноћом да С тетрамер није везан за ДНК, у ком случају се транскрипција несметано одвија до краја гена, и 2) вероватноћом да приликом транскрипције РНК полимеразе наиђе на блокаду у виду везаног С тетрамера кроз коју успева да прође са ефикасношћу $\beta < 1$ (другим речима, β је удео РНК полимеразе које прођу кроз блокаду и наставе транскрипцију у укупном броју свих РНК полимеразе које започну транскрипцију с

гена). У термодинамичком моделовању транскрипције се стандардно уводи апроксимација да је стопа синтезе транскрипта са датог промотора, тј. његова транскрипциона активност, директно пропорционална стопи иницијације транскрипције. У случају моделовања регулације Р.С за коју је карактеристично да се, осим на нивоу иницијације, одвија и на нивоу елонгације транскрипције, ефекат блокаде пута РНК полимерази ефективно смањује стопу иницијације транскрипције са Р.С (Klimuk et al., 2018). Наиме, $\beta < 1$ чини да бројилац $1 + Z_{RNAP_L} + Z_{RNAP_R} + \beta \cdot (Z_T + Z_{RNAP_R \sim T})$ разломка у Ј-ни (3.105) буде мањи од имениоца. Увођењем израза за статистичке тежине (3.100)-(3.103) у Ј-не (3.104) и (3.105) и узимањем да је $[C] \approx [Mon]$ добијене су једначине за транскрипционе активности промотора:

$$\varphi_m(C) = \alpha \frac{f}{1 + f + g + \frac{1+g}{h} [C]^4} \quad (3.106)$$

$$\varphi_c(C) = \alpha \frac{g \left(1 + \frac{1}{h} [C]^4\right) \cdot \left(1 + f + g + \beta \frac{1+g}{h} [C]^4\right)}{\left(1 + f + g + \frac{1+g}{h} [C]^4\right)^2}, \quad (3.107)$$

где је претпостављена иста стопа транскрипције по напуштању промотора, α , са јединицама [концентрација транскрипта]/[време].

3.4.2 Динамичко моделовање експресије транскрипта и протеина система Крп2I

Једначине (3.106) и (3.107) за транскрипциону активност промотора Р.М и Р.С су употребљене као улаз за динамички модел који описује експресију транскрипта гена (Ј-не (3.108)-(3.110)) и протеина система (Ј-не (3.111)-(3.113)):

$$\frac{dc(t)}{dt} = \alpha \frac{g \left(1 + \frac{1}{h} C(t)^4\right) \cdot \left(1 + f + g + \beta \frac{1+g}{h} C(t)^4\right)}{\left(1 + f + g + \frac{1+g}{h} C(t)^4\right)^2} - \lambda_t \cdot c(t) \quad (3.108)$$

$$\frac{dm(t)}{dt} = \alpha \frac{f}{1 + f + g + \frac{1+g}{h} C(t)^4} - \lambda_t \cdot m(t) \quad (3.109)$$

$$\frac{dr(t)}{dt} = \varphi_r - \lambda_t \cdot r(t) \quad (3.110)$$

$$\frac{dC(t)}{dt} = k_P \cdot c(t) - \lambda_P \cdot C(t) \quad (3.111)$$

$$\frac{dM(t)}{dt} = k_P \cdot m(t) - \lambda_P \cdot M(t) \quad (3.112)$$

$$\frac{dR(t)}{dt} = k_P \cdot r(t) - \lambda_P \cdot R(t), \quad (3.113)$$

где су претпостављене исте константе деградације транскрипта (λ_t), константе транслације (k_P) и константе деградације протеина (λ_P) за све три компоненте система. Због ове претпоставке о једнаким вредностима параметара динамичког модела, предвиђања динамике експресије протеина директно одражавају динамику транскрипционе активности промотора у

систему. Како је експресија гена за рестрикциону ендонуклеазу конститутивна, транскрипциона активност промотора P.R (φ_r) је константна (J-на (3.110)).

3.4.3 Рескалирање величина и одређивање вредности параметара модела P-M система Kpn2I

Како апсолутне вредности концентрација и параметара нису неопходне за анализу динамике система, извршено је рескалирање величина модела ради смањивања броја слободних параметара чије је вредности неопходно проценити (Klimuk et al., 2018). Обе стране једначина (3.108)-(3.113) су подељене са $\lambda_t^4 \sqrt[4]{h}$ и уведене су рескалиране променљиве $\tau = \lambda_t t$, $\bar{m} = m/\sqrt[4]{h}$, $\bar{c} = c/\sqrt[4]{h}$, $\bar{r} = r/\sqrt[4]{h}$, $\bar{M} = M/\sqrt[4]{h}$, $\bar{C} = C/\sqrt[4]{h}$ и $\bar{R} = R/\sqrt[4]{h}$, и рескалирани параметри $\bar{\alpha} = \alpha/(\lambda_t^4 \sqrt[4]{h})$, $\bar{\varphi}_r = \varphi_r/(\lambda_t^4 \sqrt[4]{h})$, $\bar{k}_p = k_p/\lambda_t$ и $\bar{\lambda}_p = \lambda_p/\lambda_t$. На тај начин је добијен коначни облик једначина модела:

$$\frac{d\bar{c}(\tau)}{d\tau} = \bar{\alpha} \frac{g(1+\bar{C}(\tau)^4) \cdot (1+f+g+\beta \cdot (1+g) \cdot \bar{C}(\tau)^4)}{(1+f+g+(1+g) \cdot \bar{C}(\tau)^4)^2} - \bar{c}(\tau) \quad (3.114)$$

$$\frac{d\bar{m}(\tau)}{d\tau} = \bar{\alpha} \frac{f}{1+f+g+(1+g) \cdot \bar{C}(\tau)^4} - \bar{m}(\tau) \quad (3.115)$$

$$\frac{d\bar{r}(\tau)}{d\tau} = \bar{\varphi}_r - \bar{r}(\tau) \quad (3.116)$$

$$\frac{d\bar{C}(\tau)}{d\tau} = \bar{k}_p \cdot \bar{c}(\tau) - \bar{\lambda}_p \cdot \bar{C}(\tau) \quad (3.117)$$

$$\frac{d\bar{M}(\tau)}{d\tau} = \bar{k}_p \cdot \bar{m}(\tau) - \bar{\lambda}_p \cdot \bar{M}(\tau) \quad (3.118)$$

$$\frac{d\bar{R}(\tau)}{d\tau} = \bar{k}_p \cdot \bar{r}(\tau) - \bar{\lambda}_p \cdot \bar{R}(\tau). \quad (3.119)$$

За процену вредности параметара модела су коришћени резултати експеримената које су наши сарадници извели на систему Kpn2I (Klimuk et al., 2018). У једном од експеримената, ДНК фрагмент који садржи одређени промотор система је клониран узводно од *galK* гена на плазмиду, тако да се експресија тог гена врши са датог промотора, *in vivo*. Притом је у неке ћелије паралелно унет и плазмид са *c* геном и његовим промотором, који омогућава експресију *C* протеина у ћелији. Ради мапирања промотора, односно одговарајућег места почетка транскрипције, на изоловане транскрипте *инициране* са испитиваног промотора је примењена метода продужења прајмера специфичних за *galK* ген. Као резултат раздвајања обележених продуката електрофорезом, добијене су траке на гелу чија позиција указује на то где се налази место почетка транскрипције са датог промотора. Како су овакве траке добијене у одсуству и у присуству *C* протеина у ћелији, за *m* транскрипт је постављен математички услов да је однос његових равнотежних концентрација (\bar{m}_{eq} у J-ни (3.115) за које важи $d\bar{m}(\tau)/d\tau = 0$) у ова два случаја једнак количнику интензитета добијених трака на гелу, што захтева да овом количнику буде једнак однос транскрипционих активности P.M у одсуству ($\bar{C}_{eq} = 0$) и у присуству *C* протеина:

$$\frac{\bar{\alpha} \cdot f}{1+f+g} \cdot \left(\frac{\bar{\alpha} \cdot f}{1+f+g+(1+g) \cdot \bar{C}_{eq}^4} \right)^{-1} = 86.229. \quad (3.120)$$

Сличан услов је постављен и за c транскрипт (Ј-на (3.114)), с тим што у овом случају количнику интензитета трака одговара однос стопа *иницијације* транскрипције са Р.С (члан који представља вероватноћу елонгације транскрипта се занемарује):

$$\frac{\bar{\alpha} \cdot g}{1 + f + g} \cdot \left(\frac{\bar{\alpha} \cdot g \cdot (1 + \bar{C}_{\text{eq}}^4)}{1 + f + g + (1 + g) \cdot \bar{C}_{\text{eq}}^4} \right)^{-1} = 0.263. \quad (3.121)$$

Ефекат блокирања пута РНК полимеразе током транскрипције c гена је испитан извођењем *in vitro* транскрипције са Р.С у присуству и у одсуству С протеина и анализом добијене количине (која је пропорционална интензитету траке на гелу) комплетних транскрипата очекиване дужине. Количник интензитета трака добијених у присуству и у одсуству С протеина одговара односу равнотежних концентрација c транскрипта (\bar{C}_{eq} у Ј-ни (3.114) за које важи $d\bar{c}(\tau)/d\tau = 0$) у овим случајевима, који је управо једнак другом члану израза за транскрипциону активност Р.С, који даје вероватноћу елонгације транскрипта (пошто се члан који описује иницијацију транскрипције пократи):

$$\frac{1 + f + g + \beta \cdot (1 + g) \cdot \bar{C}_{\text{eq}}^4}{1 + f + g + (1 + g) \cdot \bar{C}_{\text{eq}}^4} = 0.44. \quad (3.122)$$

Како су експериментални резултати указали на приближно једнаке количине транскрипата сва три гена у равнотежном стању (Klimuk et al., 2018), уведен је услов $\bar{\varphi}_m(\bar{C}_{\text{eq}}) \approx \bar{\varphi}_c(\bar{C}_{\text{eq}}) \approx \bar{\varphi}_r$ на основу којег је изведена следећа једнакост:

$$\frac{f}{g} = \frac{(1 + \bar{C}_{\text{eq}}^4) \cdot (f/(1 + g) + 1 + \beta \cdot \bar{C}_{\text{eq}}^4)}{f/(1 + g) + 1 + \bar{C}_{\text{eq}}^4}. \quad (3.123)$$

Уз стандардни услов за равнотежно стање ($d\bar{c}(\tau)/d\tau = d\bar{C}(\tau)/d\tau = 0$) примењен на Ј-не (3.114) и (3.117) из којег следи

$$\bar{\varphi}_c(\bar{C}_{\text{eq}}) = \bar{\alpha} \frac{g(1 + \bar{C}_{\text{eq}}^4) \cdot (1 + f + g + \beta \cdot (1 + g) \cdot \bar{C}_{\text{eq}}^4)}{(1 + f + g + (1 + g) \cdot \bar{C}_{\text{eq}}^4)^2} = \frac{\bar{\lambda}_p}{\bar{k}_p} \cdot \bar{C}_{\text{eq}}^4, \quad (3.124)$$

решавањем система једначина (3.120)-(3.124) једнозначно су одређене вредности параметара $f = 2.9$, $g = 0.02$, $\bar{\alpha} = 28$ и $\beta = 0.44$. Параметрима $\bar{k}_p = 3$ и $\bar{\lambda}_p = 1/6$ су додељене стандардне вредности које задају да транскрипт током свог трајања буде 3 пута преведен у протеин транслацијом и да се протеини разграђују 6 пута спорије од транскрипата (Phillips et al., 2012; Snepren & Zocchi, 2005).

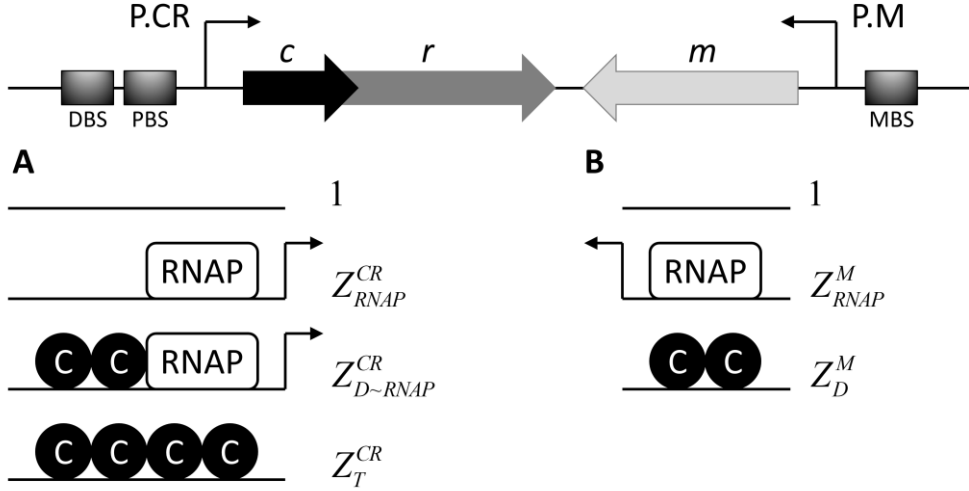
3.5 Биофизичко моделовање Р-М система Esp1396I

3.5.1 Термодинамичко моделовање регулације транскрипције гена система Esp1396I

Према сазнањима која потичу из ранијег експерименталног истраживања Р-М система Esp1396I (Bogdanova et al., 2009), његов промотор Р.СR, са којег се врши транскрипција c и r гена у оквиру оперона, је регулисан С протеином на исти начин као у систему AhdI, па у термодинамички модел улазе конфигурације (Слика 7А):

- 1) празне ДНК промотора и регулаторних секвенци (статистичка тежина 1),

- 2) РНК полимеразе везане за промотор у режиму базалне транскрипције гена (Z_{RNAP}^{CR} , Ј-на (3.12)),
- 3) С димера везаног за DBS који регрутује РНК полимеразу на промотор и активира транскрипцију ($Z_{D\sim RNAP}^{CR}$, Ј-на (3.14)), и
- 4) два С димера везана за ДНК који утишавају транскрипцију, при чему С димер везан за DBS регрутује С димер на PBS (Z_T^{CR} , Ј-на (3.15)).



Слика 7. Графички приказ регулације транскрипције у Р-М систему Esp1396I. Организација гена система на плазмиду је приказана одговарајуће именованим широком стрелицама, танке стрелице означавају смерове транскрипције са промотора Р.СР и Р.М, док правоугаоници означени са DBS, PBS и MBS представљају дистално, проксимално и место за везивање С димера у региону промотора Р.М. Илустроване су дозвољене конфигурације РНК полимеразе (правоугаоник RNAP) и С протеина (круг обележен са С представља мономер) на ДНК промотора (А) Р.СР и (Б) Р.М, а изрази за њихове назначене статистичке тежине су изведени у главном тексту. Конфигурације које садрже стрелицу су транскрипционо активне. Слика адаптирана из (Graovac et al., 2019).

Разлика у односу на термодинамички модел регулације Р.СР у систему AhdI (Ј-на (3.11)), који је представљен у поглављу 3.1.1, је у одсуству конфигурације у којој се на ДНК налази само један С димер, тј. транскрипциона активност је дата једначином:

$$\varphi_{cr} = \alpha \frac{Z_{RNAP}^{CR} + Z_{D\sim RNAP}^{CR}}{1 + Z_{RNAP}^{CR} + Z_{D\sim RNAP}^{CR} + Z_T^{CR}}. \quad (3.125)$$

Као што је раније објашњено, заступљеност ове конфигурације је изузетно ниска у поређењу са осталим наведеним због, са једне стране, малог афинитета PBS за везивање С протеина, а са друге, високе кооперативности у везивању другог С димера или РНК полимеразе када је С димер везан за DBS. Модел без ове конфигурације је експериментално проверен на примеру система AhdI (Vogdanova et al., 2008).

Еквивалентним увођењем константних параметара који апсорбују концентрацију слободне РНК полимеразе и слободне енергије протеин-протеин и протеин-ДНК интеракција (види Ј-не (3.17), (3.19) и (3.20)), добија се зависност транскрипционе активности Р.СР од концентрације мономера С протеина:

$$\varphi_{cr}(C) = \alpha \frac{a + b[C]^2}{1 + a + b[C]^2 + c[C]^4}, \quad (3.126)$$

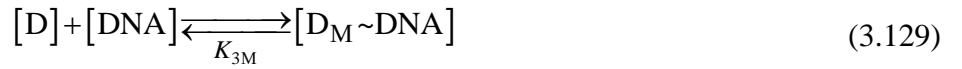
при чему је уведена апроксимација $[C] \approx [Mon]$. Параметар α у Ј-ни (3.126) је изражен преко константе $\varphi_{cr}^{basal} = \alpha \cdot a / (1 + a)$, која представља вредност базалне транскрипционе активности Р.СР када $[C] \rightarrow 0$, чиме је добијена коначна једначина модела:

$$\varphi_{cr}(C) = \varphi_{cr}^{basal} \frac{1 + \frac{b}{a}[C]^2}{1 + \frac{b}{a+1}[C]^2 + \frac{c}{a+1}[C]^4}. \quad (3.127)$$

У Р-М систему Esp1396I, С протеин регулише и промотор Р.М, па у модел морају ући његове следеће конфигурације (Слика 7В) (Morozova et al., 2016):

- 1) конфигурација (концентрације [DNA]) празне ДНК,
- 2) конфигурација (концентрације [RNAP~DNA]) РНК полимеразе везане за промотор у режиму базалне транскрипције, и
- 3) конфигурација (концентрације [D_M~DNA]) С димера везаног за МBS.

Наведене конфигурације настају у следећим равнотежним реакцијама везивања:



чије су равнотежне константе дисцијације дате изразима:

$$K_{2M} = ([RNAP][DNA]) / [RNAP \sim DNA] = k e^{\Delta G_{P.M}} \quad (3.130)$$

$$K_{3M} = ([D][DNA]) / [D_M \sim DNA] = k e^{\Delta G_{MBS}}, \quad (3.131)$$

где је ΔG_{MBS} слободна енергија интеракције С димера и ДНК везивног места у области Р.М. Изрази за статистичке тежине конфигурација (где је K_1 дато J-ном (3.6)),

$$Z_{RNAP}^M = [RNAP \sim DNA] / [DNA] = [RNAP] / K_{2M} = k^{-1} [RNAP] e^{-\Delta G_{P.M}} \quad (3.132)$$

$$Z_D^M = [D_M \sim DNA] / [DNA] = [C]^2 / (K_1 K_{3M}) = k^{-2} [C]^2 e^{-\Delta G_D - \Delta G_{MBS}}, \quad (3.133)$$

се могу упростити увођењем константних параметара $f = k^{-1} [RNAP] e^{-\Delta G_{P.M}}$ и $g = k^{-2} e^{-\Delta G_D - \Delta G_{MBS}}$, чиме једначина за транскрипциону активност Р.М

$$\varphi_m = \beta \frac{Z_{RNAP}^M}{1 + Z_{RNAP}^M + Z_D^M} \quad (3.134)$$

постаје

$$\varphi_m(C) = \beta \frac{f}{1 + f + g[C]^2}. \quad (3.135)$$

Број параметара је додатно смањен увођењем константне базалне транскрипционе активности Р.М, $\varphi_m^{basal} = \beta \cdot f / (1 + f)$, која следи из J-не (3.135) за $[C] \approx 0$, и параметра $K_m^2 = g / (f + 1)$, чиме је добијена коначна једначина,

$$\varphi_m(C) = \varphi_m^{basal} \frac{1}{1 + K_m^2 [C]^2}, \quad (3.136)$$

која је коришћена у даљем моделовању.

3.5.2 Динамичко моделовање експресије транскрипата и протеина система Esp1396I

Једначине које описују промену концентрација транскрипата и протеина система Esp1396I са временом су еквивалентне једначинама динамичког модела AhdI (поглавље 3.1.2):

$$\frac{dcr(t)}{dt} = \varphi_{cr}(C(t)) - \lambda_{cr} \cdot cr(t) \quad (3.137)$$

$$\frac{dm(t)}{dt} = \varphi_m(C(t)) - \lambda_m \cdot m(t) \quad (3.138)$$

$$\frac{dC(t)}{dt} = k_C \cdot cr(t) - \lambda_C \cdot C(t) \quad (3.139)$$

$$\frac{dR(t)}{dt} = k_R \cdot cr(t) - \lambda_R \cdot R(t) \quad (3.140)$$

$$\frac{dM(t)}{dt} = k_M \cdot m(t) - \lambda_M \cdot M(t). \quad (3.141)$$

И овде је уведен услов једнакости константи деградације сва три протеина, па је диференцијална једначина (3.140) замењена одговарајућом алгебарском,

$$R(t) = \frac{k_R}{k_C} C(t). \quad (3.142)$$

У складу са показаним деловањем спрегнуте транслације транскрипата c и r гена, које је демонстрирано ранијим равнотежним мерењем експресије протеина у систему Esp1396I (Cesnaviciene et al., 2003), уведено је да је $k_R = k_C \cdot 2/3$.

3.5.3 Одређивање вредности параметара модела Р-М система Esp1396I фитовањем података из експерименталног мерења динамике експресије протеина у ћелији

Наши сарадници су спровели експериментална мерења експресије рестрикционе ендонуклеазе и метилтрансферазе система Esp1396I у времену, *in vivo*, на нивоу појединачних ћелија (Morozova et al., 2016). За фитовање је одабран репрезентативни пар добијених трајекторија које приказују динамику експресије ових протеина у ћелији од њеног трансформисања Р-М системом на плазмиду. Паралелно је бележен и број ћелија у култури што је искоришћено за одређивање брзине деобе ћелија.

Моделовање брзине деобе ћелија. На графику зависности логаритма броја ћелија у култури од времена се разликују два режима раста у којима су се ћелије делиле приближно константним брзинама (Слика 8): бржи раст током почетних ~ 160 минута, и спорији раст у наредном периоду. До промене у брзини раста ћелија је вероватно дошло када су бактерије потрошиле приоритетни супстрат из медијума и почеле да метаболишу алтернативне нутријенте (Baev et al., 2006). У сваком од ових периода, промена броја ћелија (N) у култури се може описати експоненцијалним растом (Neuhauser, 2011):

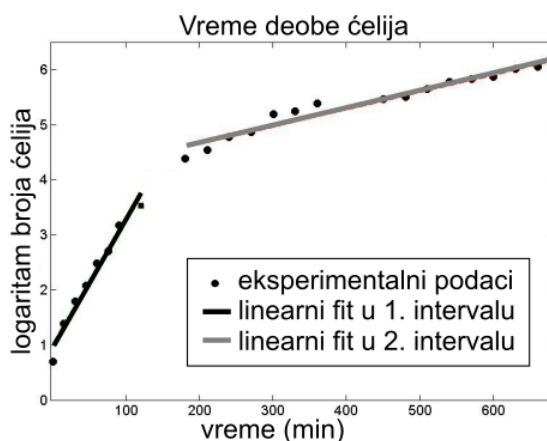
$$\frac{dN(t)}{dt} = v \cdot N(t) \Rightarrow N(t) = N_0 \cdot e^{vt} \quad (3.143)$$

где је v константна стопа раста по ћелији. Како је

$$\frac{d \ln(N(t))}{dt} = v, \quad (3.144)$$

вредности v у првом и другом временском интервалу експеримента су одређене независно из нагиба правих којима су фитовани подаци о зависности $\ln(N)$ од t . Израчунато је да је време

деобе ћелије $\tau = \ln(2)/v$ у првом интервалу експеримента износило ~ 40 минута, а у другом ~ 320 минута (Morozova et al., 2016).



Слика 8. Фитовање података о динамици раста културе у експерименту у коме је мерена динамика експресије протеина Р-М система Esp1396I. Зависност логаритма измереног броја ћелија од времена (кружићи) је фитована једначином праве у два одвојена временска интервала (тамнија и светлија пуна линија). Слика адаптирана из (Morozova et al., 2016).

Нумеричко решавање система једначина модела и фитовање података. Подаци о концентрацијама $R(t)$ и $M(t)$ у релативним јединицама су фитовани одговарајућим кривама које генерише нумерички решен модел (Morozova et al., 2016). Прецизније, за сваки параметар је дефинисан низ његових могућих вредности тако да покрије биолошки релевантан опсег (Sperren & Zocchi, 2005) и састављене су све могуће комбинације вредности параметара модела. Као почетни услов за симулацију динамике је узето да су концентрације свих транскрипта и протеина једнаке нули у тренутку првог мерења, извршеног непосредно по уносу Р-М система у ћелију, које је означено као $t = 0$. За сваку комбинацију вредности параметара су нумерички решене једначине модела, методом Runge-Kutta, и одређена је сума квадрата одступања моделом предвиђених од експериментално одређених вредности $R(t)$ и $M(t)$. Најбољи фит модела, односно одговарајућа комбинација вредности параметара, је нађена као минимум ове суме.

Овај задатак захтева рачунски захтевну претрагу у 12-димензионалном параметарском простору са потенцијално великим бројем редундантних решења. Ради брже и тачније претраге, она је изведена у два корака, независно за $R(t)$ и $M(t)$, што је могуће због одсуства утицаја експресије метилтрансферазе на експресију рестрикционе ендонуклеазе. Прво су фитовани подаци за $R(t)$ нумеричким решавањем Ј-на (3.137), (3.139) и (3.142) и на тај начин су одређене вредности 7 параметара. У другом кораку су фитовани подаци за $M(t)$ нумеричким решавањем Ј-на (3.138) и (3.141) и одређене су вредности преосталих 5 параметара. Како експресију m гена регулише C протеин, као прокси за зависност његове концентрације од времена је искоришћен веома добар фит израза $p_1 t^{p_2}$ на податке за $R(t)$, у складу са претпоставком модела о пропорционалности концентрација рестрикционе ендонуклеазе и C протеина (Ј-на (3.142)). На тај начин је фитовање $M(t)$ раздвојено од фитовања $R(t)$ и избегнута је истовремена претрага свих 12 параметара модела. На крају је изведено прецизно подешавање параметара динамике метилтрансферазе како би се неутралисала мала разлика између фитоване криве $R(t)$ коју предвиђа модел и коришћеног проксија.

Како су се током експеримента транскрипти и протеини губили из ћелије не само деградацијом, већ и разблаживањем услед деобе ћелије, приликом нумеричког решавања је стопа деградације сваког транскрипта/протеина замењивана брзином деобе ћелије, v , када је њена вредност била нижа од вредности v . Другим речима, претпостављено је да се у случају ниже стопе деградације молекули доминантно елиминишу деобом ћелије. Вредности

параметара модела за које је добијен најбољи фит експерименталних података су наведене у Табели 2.

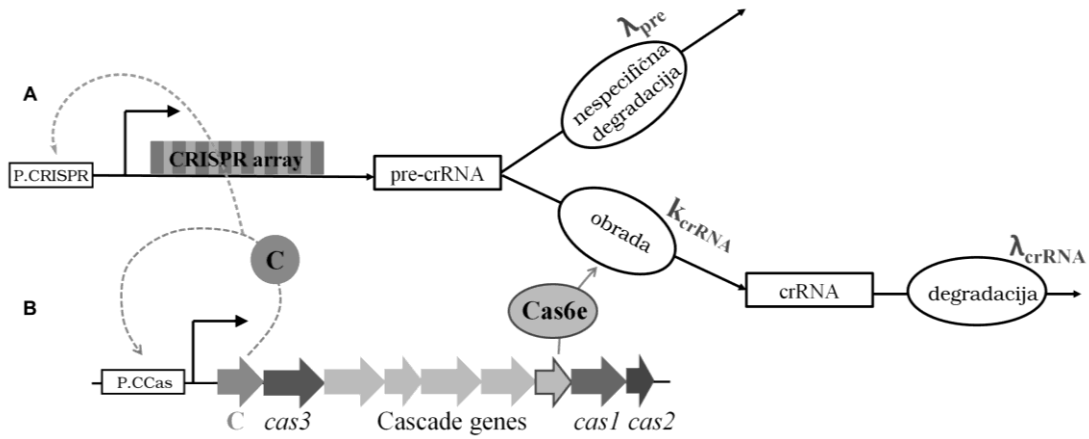
Табела 2. Вредности параметара модела Р-М система Esp1396I.

Константа	Опис	Вредност
φ_{cr}^{basal}	Базална транскрипциона активност Р.СР	0.25
φ_m^{basal}	Базална транскрипциона активност Р.М	36
a		0.4
b	Константе које апсорбују релевантне слободне енергије	$2.1 \cdot 10^{-7}$
c	протеин-протеин и протеин-ДНК интеракција и	$1.0 \cdot 10^{-16}$
K_m	концентрацију слободне РНК полимеразе у ћелији	$7.5 \cdot 10^{-2}$
k_C	Константа транслације за С протеин	0.9
k_R	Константа транслације за рестрикциону ендонуклеазу	0.6
k_M	Константа транслације за метилтрансферазу	0.55
λ_{cr}	Константа деградације <i>cr</i> транскрипта	$6.5 \cdot 10^{-3}$
λ_m	Константа деградације <i>m</i> транскрипта	$2.5 \cdot 10^{-3}$
λ_C	Константа деградације С протеина	$2.5 \cdot 10^{-3}$
λ_R	Константа деградације рестрикционе ендонуклеазе	$2.5 \cdot 10^{-3}$
λ_M	Константа деградације метилтрансферазе	$2.5 \cdot 10^{-3}$

3.6 Биофизичко моделовање система CRISPR-Cas

Предложена генска кола за регулацију експресије *cr*RNA подразумевају да се у ћелију која поседује CRISPR низ под контролом промотора Р.CRISPR (Слика 9А) помоћу вектора (нпр. плазида) унесу *cas* гени са придруженим геном за С протеин Р-М система AhdI под контролом заједничког промотора названог Р.CCas (Слика 9В) (Rodic, Blagojevic, Djordjevic, et al., 2017). Моделована су три таква генска кола која се разликују по механизму регулације промотора Р.CCas:

- 1) „**Гранични модел**“ описује максималну теоријску брзину индуковања експресије *cr*RNA и обезбеђује визуелизацију граничне динамике система у односу на коју се могу интерпретирати резултати симулације остала два, биолошки реалистичнија модела. У овом моделу, стопа обраде *pre-cr*RNA у $t = 0$ тренутно порасте са почетне на задату равнотежну вредност.
- 2) „**Конститутивни модел**“ подразумева постепену експресију *c* и *cas* гена са конститутивног (базално активног) промотора константне транскрипционе активности.
- 3) „**Кооперативни модел**“ представља случај у коме су *c* и *cas* гени стављени под контролу Р.СР промотора из Р-М система AhdI који је регулисан С протеином. У зависности од конфигурације његових везаних молекула за ДНК, С протеин делује као активатор или као репресор транскрипције (видети поглавље 3.1.1).



Слика 9. Основна шема предложених генских кола за покретање регулисане експресије crRNK. (A) Pre-crRNK (чија је количина означена правоугаоником), као што стрелице показују, настаје транскрипцијом CRISPR низа са промотора P.CRISPR, а губи се неспецифичном деградацијом или обрадом у crRNK коју спроводи протеин Cas6e (обојени овал). CrRNK се елиминише разградњом. Уз биохемијске процесе означене овалима су наведене њихове константе брзине, при чему је k_{crRNA} сразмерна концентрацији Cas6e. (B) На плазмиду који се уводи у ћелију, која већ поседује CRISPR низ, се налазе гени за C и Cas протеине под контролом P.Ccas промотора. У зависности од испитиваног сценарија, P.CCas и/или P.CRISPR могу бити регулисани C протеином, што је илустровано испрекиданим стрелицама. Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

Ефекти постепене и регулисане синтезе Cas протеина, као и брзе неспецифичне деградације pre-crRNK су испитани *in silico* коришћењем модела предложених генских кола уз претпоставку о константној транскрипционој активности промотора P.CRISPR. За потребе испитивања значаја заједничке активације експресије cas гена и CRISPR низа, до које највероватније долази у *wt* CRISPR-Cas систему, предложени су и механизми регулације P.CRISPR и уведени су у одговарајуће моделе генских кола.

3.6.1 Динамичко моделовање експресије RNK и протеина синтетичких генских кола базираних на систему CRISPR-Cas

Динамичко моделовање експресије pre-crRNK и crRNK. Динамички модел експресије pre-crRNK и crRNK дат једначинама:

$$\frac{dpre-crRNA(t)}{dt} = \varphi_{crispr} - (\lambda_{pre} + k_{crRNA}) \cdot pre-crRNA(t) \quad (3.145)$$

$$\frac{dcrRNA(t)}{dt} = k_{crRNA} \cdot pre-crRNA(t) - \lambda_{crRNA} \cdot crRNA(t), \quad (3.146)$$

је раније постављен (Djordjevic et al., 2012) и подразумева да се (Слика 9A): 1) транскрипт CRISPR низа (концентрације pre-crRNA) синтетише са промотора P.CRISPR транскрипционе активности φ_{crispr} , неспецифично разграђује са стопом деградације λ_{pre} , а у crRNK преводи специфичном обрадом од стране протеина Cas6e са стопом k_{crRNA} ; 2) кратке, регулаторне crRNK (концентрације crRNA) производе специфичним сечењем дугих pre-crRNK, а разграђују са стопом деградације λ_{crRNA} .

Стопа обраде pre-crRNK зависи од концентрације кључног протеина Cas6e који врши ову обраду, према једначини:

$$k_{crRNA}(t) = \tilde{k}_{crRNA} \cdot Cas6e(t), \quad (3.147)$$

где је \tilde{k}_{crRNA} константа обраде. Pre-crRNK су, пре свега због веома брзе неспецифичне деградације, у ћелији присутне у малом броју (<10 молекула према резултатима мерења у (Rougach et al., 2010)), па се може претпоставити линеарна зависност брзине обраде супстрата од стране Cas протеина који су у вишку. Промотори P.CRISPR и P.Cas у *wt* систему CRISPR-Cas су довољно јаки да се детерминистичко решење једначина поклапа са усредњеним стохастичким (Djordjevic et al., 2012), па је и у овој анализи задржан детерминистички модел који омогућава знатно једноставнију анализу општих (квалитативних) особина понашања система у времену.

Динамичко моделовање експресије протеина C и Cas6e. Следеће једначине динамичког модела описују промене концентрација заједничког транскрипта *c* и *cas* гена (*ccas*), као и C и Cas6e протеина са временом:

$$\frac{d\text{ccas}(t)}{dt} = \varphi_{ccas} - \lambda_{ccas} \cdot \text{ccas}(t) \quad (3.148)$$

$$\frac{dC(t)}{dt} = k_C \cdot \text{ccas}(t) - \lambda_C \cdot C(t) \quad (3.149)$$

$$\frac{d\text{Cas6e}(t)}{dt} = k_{\text{Cas6e}} \cdot \text{ccas}(t) - \lambda_{\text{Cas6e}} \cdot \text{Cas6e}(t), \quad (3.150)$$

где су λ_{ccas} , λ_C и λ_{Cas6e} константе деградације *ccas* транскрипта, односно C и Cas6e протеина, а k_C и k_{Cas6e} константе транслације за C и Cas6e протеине. Уз претпоставку сличних стопа деградације протеина ($\lambda_C = \lambda_{\text{Cas6e}}$), диференцијална једначина (3.150) се може заменити алгебарском,

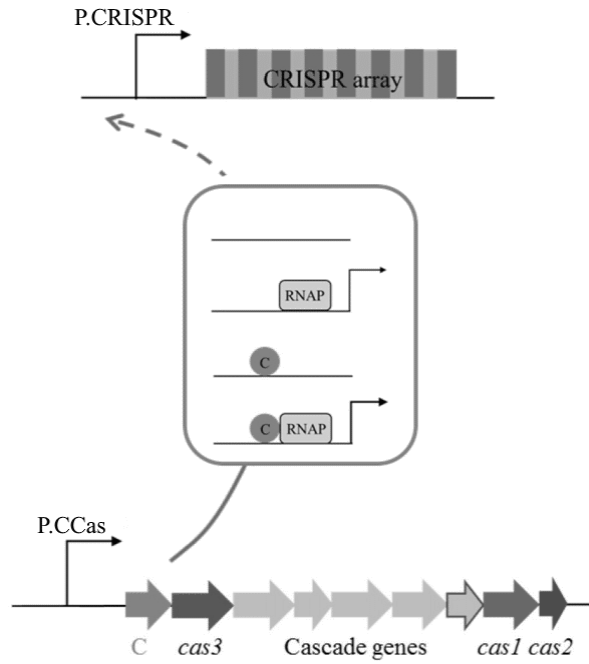
$$\text{Cas6e}(t) = \frac{k_{\text{Cas6e}}}{k_C} C(t). \quad (3.151)$$

3.6.2 Термодинамичко моделовање регулације транскрипције у „конститутивном моделу“ експресије система CRISPR-Cas

У „конститутивном моделу“, промотор P.CCas се одликује константном транскрипционом активношћу (φ_{ccas}). У случају када се *не* подразумева константна транскрипциона активност P.CRISPR (φ_{crispr}), претпоставља се да по уласку плаزمида у ћелију постепено синтетисани C протеин почиње да активира транскрипцију са P.CRISPR најједноставнијим могућим механизмом, у коме C мономер везан за ДНК регрутује РНК полимеразу на промотор (Rodić, Vladojevic, Djordjevic, et al., 2017).

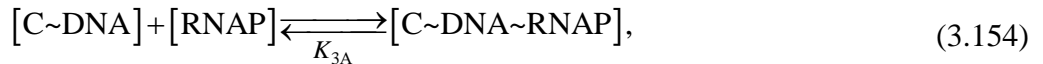
Промотор P.CRISPR тако регулисан C протеином се може наћи у следећим конфигурацијама (Слика 10):

- 1) празна ДНК промоторског региона (концентрације [DNA]),
- 2) РНК полимеразе везана за промотор ([RNAP~DNA]) који је, услед тога, базално транскрипционо активан,
- 3) један мономер C протеина везан за специфично везивно место у близини промотора ([C~DNA]),
- 4) РНК полимеразе регрутована на промотор од стране везаног C протеина ([C~DNA~RNAP]) који тиме активира транскрипцију CRISPR низа (претпостављено је да је интеракција C протеина са РНК полимеразом умерене јачине што повећава вероватноћу конфигурације самог C протеина на ДНК и оправдава њено укључивање у модел).



Слика 10. Шематски приказ регулације транскрипције у „конститутивном моделу“. Транскрипциона активност P.CCas је, у овом моделу увек константна, док транскрипциона активност P.CRISPR, у зависности од испитиваног сценарија, може бити или константна, или регулисана C протеином (испрекидана сива стрелица) на начин описан уоквиреним дозвољеним конфигурацијама. Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

Следеће хемијске реакције са одговарајућим равнотежним константама дисоцијације (K) описују формирање ових конфигурација:



па су статистичке тежине конфигурација дате следећим изразима:

$$Z_{\text{RNAP}}^{\text{CRISPR}} = \frac{[\text{RNAP} \sim \text{DNA}]}{[\text{DNA}]} = \frac{[\text{RNAP}]}{K_{1A}} \quad (3.155)$$

$$Z_{\text{C}}^{\text{CRISPR}} = \frac{[\text{C} \sim \text{DNA}]}{[\text{DNA}]} = \frac{[\text{C}]}{K_{2A}} \quad (3.156)$$

$$Z_{\text{C} \sim \text{RNAP}}^{\text{CRISPR}} = \frac{[\text{C} \sim \text{DNA} \sim \text{RNAP}]}{[\text{DNA}]} = \frac{[\text{RNAP}][\text{C}]}{K_{2A} K_{3A}}. \quad (3.157)$$

Увођењем параметара $d = [\text{RNAP}]/K_{1A}$, $e = 1/K_{2A}$ и $f = K_{1A}/K_{3A}$ у израз за транскрипциону активност P.CRISPR,

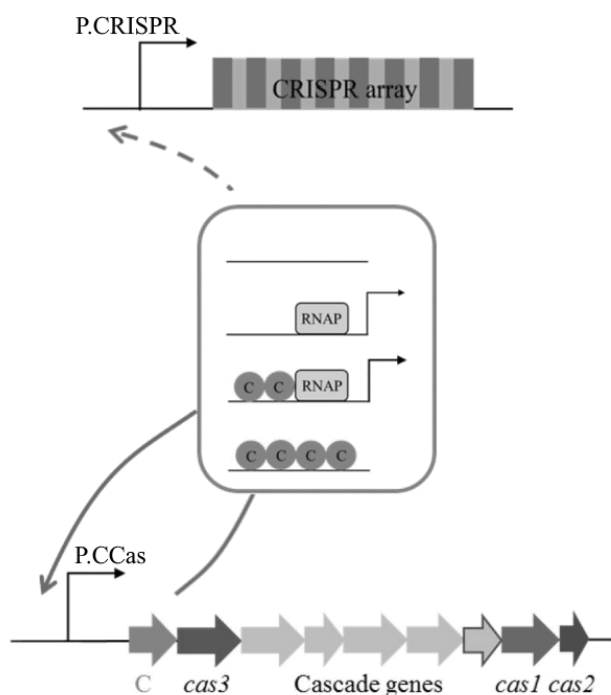
$$\varphi_{\text{crispr}} = \gamma \frac{Z_{\text{RNAP}}^{\text{CRISPR}} + Z_{\text{C} \sim \text{RNAP}}^{\text{CRISPR}}}{1 + Z_{\text{RNAP}}^{\text{CRISPR}} + Z_{\text{C}}^{\text{CRISPR}} + Z_{\text{C} \sim \text{RNAP}}^{\text{CRISPR}}} \quad (3.158)$$

изведен у складу са стандардном претпоставком према којој је она пропорционална равнотежној вероватноћи везивања РНК полимеразе за промотор (Shea & Ackers, 1985) са константом пропорционалности γ , добијена је коначна једначина:

$$\varphi_{crispr}(C) = \gamma \frac{d + d \cdot e \cdot f[C]}{1 + d + e[C] + d \cdot e \cdot f[C]} \quad (3.159)$$

3.6.3 Термодинамичко моделовање регулације транскрипције у „кооперативном моделу“ експресије система CRISPR-Cas

У „кооперативном моделу“ су оба промотора – или само P.CCas када је φ_{crispr} константа – регулисана подробно проученим механизмом регулације промотора P.CR у P-M систему AhdI (Bogdanova et al., 2008; Rodic, Blagojevic, Djordjevic, et al., 2017). Као што је детаљно објашњено у поглављу 3.1.1, овај механизам подразумева да транскрипцију са релативно слабог промотора P.CR активира везивање димера C протеина за DBS, који олакшава везивање РНК полимеразе за ДНК. Међутим, исти везани C димер олакшава и везивање другог C димера за PBS чиме се транскрипција утишава, па активност промотора у сваком тренутку зависи од односа вероватноћа формирања активаторске и репресорске конфигурације по везивању C димера за DBS. Како се експресија *c* гена врши са ауторегулисаног P.CR, по уласку плаزمида који носи овај ген почиње његова слаба, базална транскрипција, да би ниво експресије нагло порастао када се синтетиче довољна количина C протеина за активацију транскрипције, а затим опао када концентрација C протеина толико порасте да репресорска конфигурација C тетрамера на ДНК постане преовлађујућа.



Слика 11. Шематски приказ регулације транскрипције у „кооперативном моделу“. Транскрипциона активност P.CCas је, у овом моделу увек регулисана C протеином (пуна сива стрелица), док транскрипциона активност P.CRISPR, у зависности од испитиваног сценарија, може бити или константна, или такође регулисана C протеином (испрекидана сива стрелица). Регулација оба промотора подразумева дозвољене конфигурације уоквирене у средишњем делу слике. Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

Моделовањем оваквог система се долази до израза за транскрипциону активност P.CR (J-на (3.16) у поглављу 3.1.1, али према оригиналном извођењу у (Bogdanova et al., 2008), без

укључивања конфигурације С димера самог на ДНК) који је директно преузет да у „кооперативном моделу“ опише транскрипцију са P.CCas:

$$\varphi_{ccas}(C) = \alpha \frac{a+b[C]^2}{1+a+b[C]^2+c[C]^4}. \quad (3.160)$$

Део параметара једначине је редефинисан како би вредности параметара могле конзистентно да се подесе и варирају у сва три модела. Конкретно, дефинисани су нови параметри тако да је $b = a \cdot p$ (видети J-ну (3.19)) и $c = p^2 \cdot q$ (видети J-ну (3.20)):

$$p = \frac{K_2}{K_1 K_3 K_5} \quad (3.161)$$

$$q = \frac{K_3 K_5^2}{K_2^2 K_4} \quad (3.162)$$

са којима једначина за транскрипциону активност P.CCas има следећи облик:

$$\varphi_{ccas}(C) = \alpha \frac{a+a \cdot p[C]^2}{1+a+a \cdot p[C]^2+p^2 \cdot q[C]^4}. \quad (3.163)$$

Како је претпостављено да је на исти начин регулисан и промотор P.CRISPR у „кооперативном моделу“, његова транскрипциона активност је дата једначином

$$\varphi_{crispr}(C) = \hat{\alpha} \frac{\hat{a} + \hat{a} \cdot \hat{p}[C]^2}{1 + \hat{a} + \hat{a} \cdot \hat{p}[C]^2 + \hat{p}^2 \cdot \hat{q}[C]^4}, \quad (3.164)$$

где су са $\hat{\alpha}$, \hat{a} , \hat{p} и \hat{q} означени параметри по аналогији са одговарајућим параметрима у једначини (3.163).

3.6.4 Одређивање и подешавање вредности параметара модела система CRISPR-Cas

Претпоставка о константној вредности φ_{crispr} . У анализи се полази од вредности параметара које су преузете из претходних радова у којима су одређене на основу резултата експеримената (Табела 3) (Bogdanova et al., 2008; Djordjevic et al., 2012). Прво је симулирана и анализирана динамика експресије сва три генска кола уз претпоставку о конститутивном P.CRISPR. Притом су задаване комбинације вредности кључних константних параметара од интереса, λ_{pre} и φ_{crispr} , и услова о равнотежном нивоу који стопа обраде, k_{crRNA} , треба да достигне. Из услова за равнотежно стање примењеног на J-ну (3.145) следи почетни услов $pre-crRNA_0 = \varphi_{crispr} / \lambda_{pre}$, док су концентрације свих осталих компонентни система на почетку 0. Жељена вредност k_{crRNA}^{eq} , изражена у јединицама $wt \lambda_{pre}$ (које износи 1 1/min), је добијана подешавањем вредности φ_{ccas} у „конститутивном моделу“, односно α (J-на (3.163)) у „кооперативном моделу“.

Табела 3. Полазне вредности параметара модела са конститутивним P.CRISPR.

Константа	Опис	Вредност	
Динамички модел експресије pre-crRNK и crRNK			
φ_{crispr}	Транскрипциона активност P.CRISPR	10	nM/min
\tilde{k}_{crRNA}	Константа обраде pre-crRNK	$2.0 \cdot 10^{-2}$	$(\text{nM} \cdot \text{min})^{-1}$
λ_{pre}	Константа деградације pre-crRNK	1.0	1/min
λ_{crRNA}	Константа деградације crRNK	$1.0 \cdot 10^{-2}$	1/min
Динамички модел експресије протеина C и Cas6e			
k_C	Константа транслације за C протеин	$6.0 \cdot 10^{-1}$	1/min
k_{Cas6e}	Константа транслације за Cas6e протеин	3.0	1/min
λ_{ccas}	Константа деградације ccas транскрипта	$2.0 \cdot 10^{-1}$	1/min
λ_C	Константа деградације C протеина	$3.3 \cdot 10^{-2}$	1/min
λ_{Cas6e}	Константа деградације Cas6e протеина	$3.3 \cdot 10^{-2}$	1/min
„Кооперативни“ термодинамички модел регулације P.CCas			
α	Константа пропорционалности за P.CCas (тј. P.CR)	1.7	nM/min
a	Константе које апсорбују слободне енергије	$1.6 \cdot 10^{-1}$	-
p	протеин-протеин и протеин-ДНК интеракција и концентрацију слободне РНК полимеразе у	$2.3 \cdot 10^{-1}$	nM ⁻²
q	ћелији	$1.5 \cdot 10^{-1}$	-

Претпоставка о временски зависној вредности φ_{crispr} . Затим је уведено да се, по уласку плаزمид са c и cas генима у ћелију, φ_{crispr} постепено промени од почетне вредности 10 nM/min до равнотежне 100 nM/min. Ту постепену промену омогућавају претпостављени механизми регулације промотора P.CRISPR описани у поглављима 3.6.2 и 3.6.3. Као и раније, симулирана је динамика система за жељене вредности равнотежне стопе обраде pre-crRNK, k_{crRNA}^{eq} , изражене у јединицама $wt \lambda_{pre}$ ($= 1 \text{ 1/min}$). Затим је симулирана динамика при различитим вредностима λ_{pre} уз фиксирану равнотежну вредност стопе обраде, $k_{crRNA}^{eq} = 100 \text{ nM/min}$.

Да би се постигао жељени скок вредности φ_{crispr} у „конститутивном моделу“, како разломак у једначини (3.159) мора бити мањи од 1, d мора бити $< 1/9$. Вредности параметара d и γ су одабране тако да задовоље овај услов. Даље, може се извести следећи израз:

$$e = \frac{99}{[C]_{eq} \cdot (f - 100)}, \quad (3.165)$$

где је $[C]_{eq}$ концентрација коју C протеин достиже у стационарном стању при подешеним вредностима параметара модела. Вредност f је фиксирана, док је вредност e рачуната из Ј-не (3.165) за $[C]_{eq}$ које одговара жељеном k_{crRNA}^{eq} .

Параметри „кооперативног модела“ (Ј-на (3.164)) су подешени на сличан начин како би се вредност φ_{crispr} променила од задате почетне до задате крајње, што води ка $\hat{a} < 1/9$. Затим је изведена једначина за $\varphi_{crispr}(C_{eq}) = 100 \text{ nM/min}$ у којој је \hat{p} изражено преко \hat{q} и на тај начин је добијен услов $\hat{q} < 1/(400 \cdot 99)$ који омогућава да се одабере и фиксира вредност \hat{q} . Вредност \hat{p} је подешавана у складу са k_{crRNA}^{eq} (тј. $[C]_{eq}$). Добијене вредности параметара термодинамичких модела регулације P.CRISPR промотора у „конститутивном“ и „кооперативном“ генском колу су наведене у Табели 4.

Табела 4. Фиксиране вредности параметара модела са регулисаним P.CRISPR.

Константа	Опис	Вредност	
„Конститутивни“ термодинамички модел регулације P.CRISPR			
γ	Константа пропорционалности за P.CRISPR	110	nM/min
d	Константе које апсорбују слободне енергије протеин-протеин и протеин-ДНК интеракција и концентрацију слободне РНК полимеразе у ћелији	$1.0 \cdot 10^{-1}$	-
e		-	nM ⁻¹
f		$2.0 \cdot 10^2$	-
„Кооперативни“ термодинамички модел регулације P.CRISPR			
$\hat{\alpha}$	Константа пропорционалности за P.CRISPR	110	nM/min
\hat{a}	Константе које апсорбују слободне енергије протеин-протеин и протеин-ДНК интеракција и концентрацију слободне РНК полимеразе у ћелији	$1.0 \cdot 10^{-1}$	-
\hat{p}		-	nM ⁻²
\hat{q}		$2.5 \cdot 10^{-5}$	-

4 Резултати

4.1 Ефекти ремећења специфичних конструктивних карактеристика Р-М система AhdI на својства динамике његове експресије

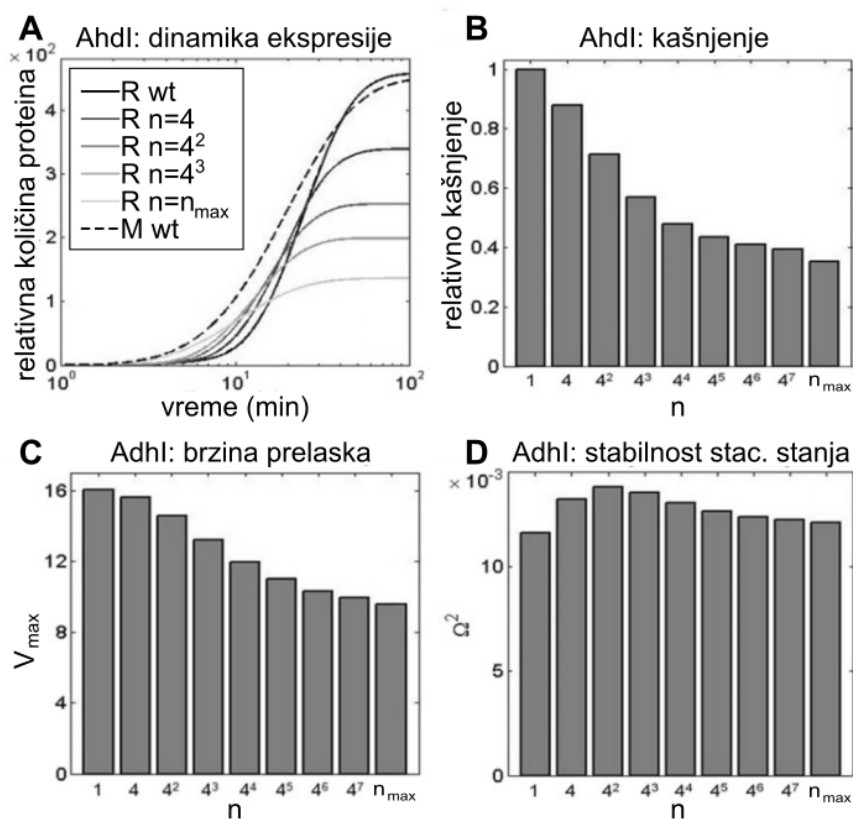
Конструктивне карактеристике специфичне за Р-М систем AhdI – претежно мономерна форма С протеина у раствору, мање ефикасна транслација c транскрипта у односу на r и m , и високо кооперативно везивање два С димера за ДНК – су постепено *in silico* неутралисане и посматрано је како се притом мења динамика експресије протеина система (Rodic, Blagojevic, Zdobnov, et al., 2017). Конкретно, мерен је степен испољености предложених значајних својстава динамике успостављања Р-М система у новој бактеријској ћелији одређивањем вредности дефинисаних квантитативних величина (поглавље 3.3.1):

- 1) τ_{delay} , које показује колико је кашњење експресије рестрикционе ендонуклеазе у мутираном у односу на кашњење у *wt* систему,
- 2) v_{max} , које представља максималан нагиб криве $R(t)$, односно указује на брзину којом експресија рестрикционе ендонуклеазе прелази из искљученог у укључено стање, и
- 3) Ω^2 , које расте са порастом стабилности стационарног стања система.

4.1.1 Повећавање афинитета С протеина за формирање хомодимера

Вредност равнотежне константе дисоцијације за реакцију димеризације С протеина је постепено смањивана од $K_1^{max} \approx 4000$ nM, која је процењена за *wt* AhdI чија је концентрација С протеина у раствору апроксимирана концентрацијом мономера, до $K_1 = 1$, што одговара случају када су сви С протеини у форми димера. Прецизније, у моделу система AhdI је варирано $K_1 = K_1^{max}/n$, где $n \in \{1, 4, 4^2, 4^3, 4^4, 4^5, 4^6, 4^7, n_{max}\}$ како би криве биле равномерно распоређене између граничних случајева, као што је објашњено у поглављу Метода 3.3.2.

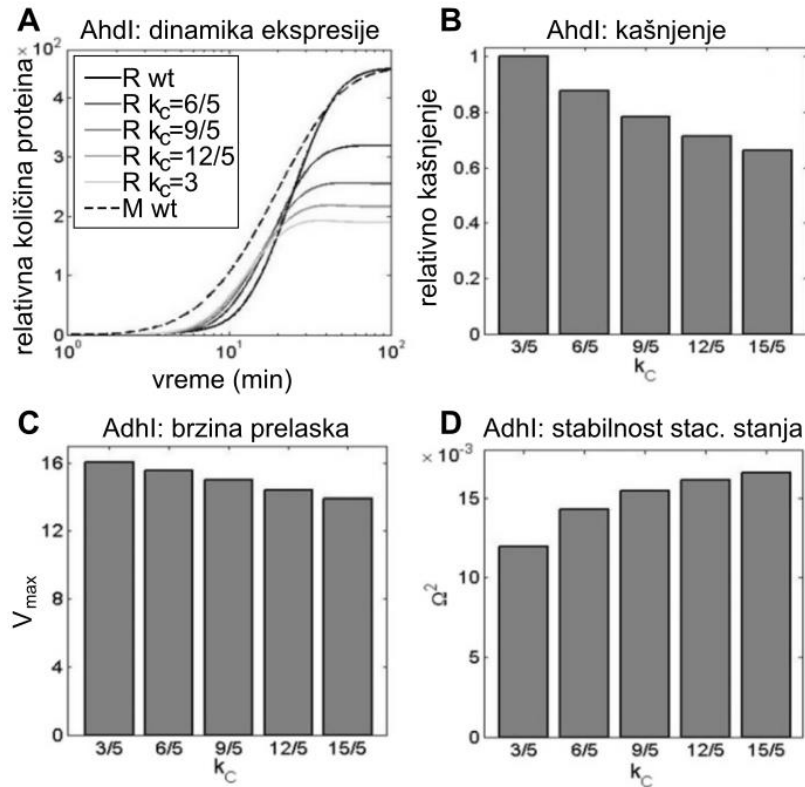
Будући да је експресија метилтрансферазе независна од експресије рестрикционе ендонуклеазе и С протеина у овом систему, и да С протеин не регулише транскрипцију са Р.М, пертурбисање K_1 не доводи до промене изгледа $M(t)$. С друге стране, смањивање K_1 значајно утиче на динамику рестрикционе ендонуклеазе, $R(t)$, на три начина (Слика 12А): 1) смањује се кашњење, тј. удаљеност почетног дела криве у односу на $M(t)$, 2) смањује се нагиб средњег дела криве, и 3) смањује се концентрација рестрикционе ендонуклеазе у стационарном стању. Ефекат смањивања K_1 на кашњење експресије рестрикционе ендонуклеазе је велики будући да релативно кашњење (τ_{delay}) са померањем равнотеже од мономера ка димерима опадне за >60% (Слика 12В). Такође, брзина преласка $R(t)$ из искљученог у укључено стање опадне за ~37% (Слика 12С), док варирање K_1 не изазива велике промене у стабилности стационарног стања (Слика 12D). Према томе, неутралисање ниског афинитета С протеина за формирање димера у систему AhdI ремети два од предложена три пожељна својства динамике система.



Слика 12. Утицај смањивања равнотежне константе дисоцијације за реакцију димеризације C протеина на (A) динамику система Ahdl, (B) релативно кашњење експресије рестрикционе ендонуклеазе, (C) брзину преласка експресије рестрикционе ендонуклеазе у укључено стање и (D) стабилност стационарног стања. Слика адаптирана из (Rodic, Blagojevic, Zdobnov, et al., 2017).

4.1.2 Повећавање ефикасности транслације транскрипта *ahdI*C

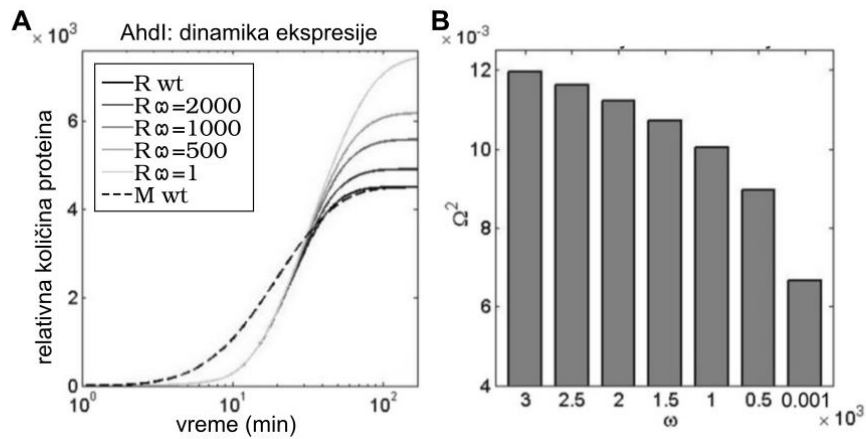
Ниска константа транслације c транскрипта, фиксирана на $3/5$ $1/\text{min}$ за *wt* Ahdl, је постепено повећавана у моделу овог система до вредности додељене константама транслације r и m транскрипата, 3 $1/\text{min}$ (Rodic, Blagojevic, Zdobnov, et al., 2017). Ефекат ове пертурбације на динамику протеина (Слика 13A) је квалитативно сличан ефекту смањивања односа концентрација мономера и димера C протеина у раствору (Слика 12A). Повећањем k_C на вредности k_R и k_M , релативно кашњење експресије рестрикционе ендонуклеазе се значајно смањи (Слика 13B, τ_{delay} опадне за $>30\%$), док v_{max} опадне за само $\sim 12\%$ (Слика 13C). Ни овај параметар нема велики утицај на стабилност стационарног стања (Слика 13D), па се може закључити да његово пертурбисање значајно ремети једно од три динамичка својства од интереса – почетно кашњење у експресији рестрикционе ендонуклеазе.



Слика 13. Утицај повећавања константе транслације за С протеин на (А) динамику система Ahdl, (В) релативно кашњење експресије рестрикционе ендонуклеазе, (С) брзину преласка експресије рестрикционе ендонуклеазе у укључено стање и (Д) стабилност стационарног стања. Слика адаптирана из (Rodic, Blagojevic, Zdobnov, et al., 2017).

4.1.3 Смањивање афинитета за кооперативно везивање два димера С протеина

Висок степен кооперативности, односно јачина интеракције два С димера на ДНК, $\omega \approx 3000$, карактеристичан за *wt* систем Ahdl, је постепено смањиван у моделу до потпуног неутралисања ове конструктивне карактеристике система, тј. $\omega \approx 1$ (Rodic, Blagojevic, Zdobnov, et al., 2017), као што је описано у поглављу Метода 3.3.3. За разлику од друге две пертурбације система, ова нема никакав утицај на кашњење експресије рестрикционе ендонуклеазе, нити на брзину преласка ове експресије у укључено стање (Слика 14А). Такође, док се променом *wt* вредности K_I и k_C концентрација рестрикционе ендонуклеазе у стационарном стању смањује, пертурбисањем ω се постиже супротан ефекат, тј. она се повећава. Варирање овог параметра има и највећи ефекат на стабилност стационарног стања (Слика 14В): у одсуству кооперативности Ω^2 има за >40% нижу вредност него у *wt* систему.

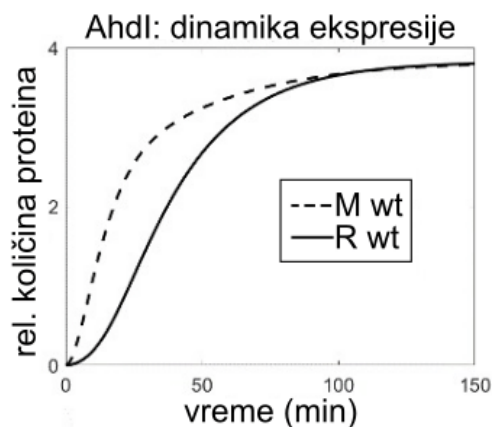


Слика 14. Утицај смањивања кооперативности у везивању С димера за ДНК на (А) динамику система AhdI и (Б) стабилност стационарног стања. Слика адаптирана из (Rodic, Blagojevic, Zdobnov, et al., 2017).

4.2 Предвиђање динамике експресије Р-М система EcoRV

Моделовањем регулације EcoRV по први пут је омогућено предвиђање динамике успостављања једног Р-М система са дивергентном архитектуром, по уласку система на плазмиду у наивну бактеријску ћелију. Модел (поглавље 3.2) описује сложену регулацију транскрипције са преклопљених промотора у којој учествује С протеин као транскрипциони фактор и узима у обзир три истакнуте конструктивне карактеристике система EcoRV: 1) доминантну форму димера С протеина у раствору, 2) стандардно високу константу транслације s транскрипта, и 3) одсуство кооперативности у везивању С димера за ДНК.

Упркос томе што из система EcoRV одсуствују карактеристике које одликују систем AhdI, динамика протеина у *wt* систему EcoRV (Слика 15) је квалитативно веома слична динамици система AhdI (Слика 5А) и поседује сва три жељена својства (Rodic, Blagojevic, Zdobnov, et al., 2017). Штавише, кашњење у експресији рестрикционе ендонуклеазе у односу на експресију метилтрансферазе је веће у систему EcoRV. Затим, $R(t)$ крива је сигмоидна, са јасним, брзим преласком из стања са ниским у стање са високим концентрацијама рестрикционе ендонуклеазе. Такође, систем достиже стабилно стационарно стање чије је $\Omega^2 > 0$ сличне вредности као у *wt* систему AhdI (види се поређењем првих стубића графика на Сликама 16С и 14В).



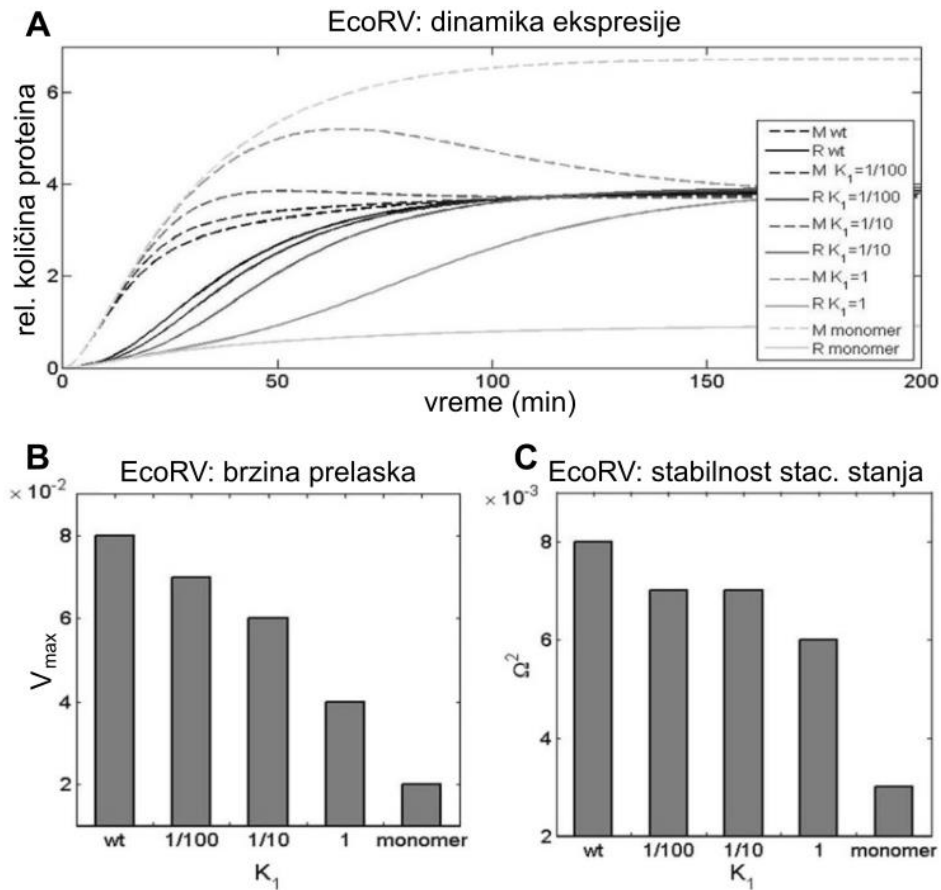
Слика 15. Динамика *wild type* система EcoRV коју предвиђа модел. Слика адаптирана из (Rodic, Blagojevic, Zdobnov, et al., 2017).

4.3 Ефекти увођења специфичних конструктивних карактеристика Р-М система AhdI у Р-М систем EсоRV на својства динамике његове експресије

Будући да се систем EсоRV одликује супротним конструктивним карактеристикама у односу на AhdI, неутралисање његових карактеристика је еквивалентно увођењу особина карактеристичних за систем AhdI (Rodic, Blagojevic, Zdobnov, et al., 2017). Конкретно, кључне карактеристике *wt* система EсоRV су *in silico* пертурбисане у смеру: 1) померања равнотеже реакције димеризације С протеина према форми мономера, 2) смањења ефикасности транслације *s* у односу на остале транскрипте система, и 3) повећања кооперативности у везивању два С димера за ДНК. Притом је, као и приликом испитивања система AhdI, праћен ефекат пертурбација на вредности три квантитативне мере испољености жељених својстава динамике експресије, τ_{delay} , v_{max} и Ω^2 .

4.3.1 Смањивање афинитета С протеина за формирање хомодимера

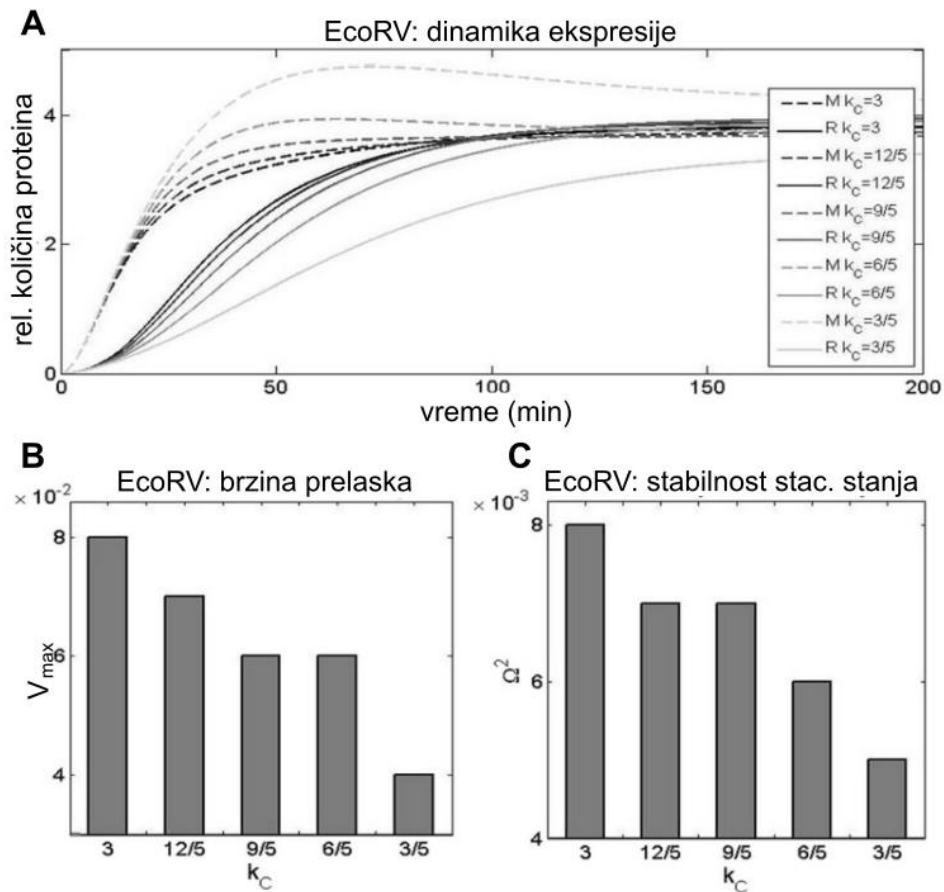
Рескалирана константа дисоцијације за реакцију димеризације С протеина, $\overline{K_1}$ (видети поглавље Метода 3.3.2), је постепено повећавана од процењене *wt* вредности (1/1000), при којој су сви С протеини у раствору у форми димера, до вредности при којој су сви С протеини мономери (10). Како у систему EсоRV постоји међузависна регулација промотора Р.СР и Р.М услед њиховог преклапања, свака пертурбација која утиче на транскрипциону активност једног ће неизбежно изменити и транскрипциону активност другог промотора, што је видљиво и у овом случају (Слика 16А). Повећање $\overline{K_1}$ не мења динамику система током првих 10 минута од уноса система у ћелију-домаћина, односно релативно кашњење је исто у пертурбисаним и у *wt* систему, али зато видно мења динамику у каснијем временском интервалу. Преласком из граничног случаја у коме су само С димери у раствору у гранични случај у коме су сви С протеини мономери се значајно измене друга два посматрана својства динамике: брзина раста експресије рестрикционе ендонуклеазе опадне за ~75% (Слика 16В), а стабилност стационарног стања система за >60% (Слика 16С). Дакле, пертурбисање ниске константе дисоцијације за реакцију димеризације С протеина у систему EсоRV има негативан утицај на два од три пожељна својства динамике експресије.



Слика 16. Утицај повећавања равнотежне константе дисоцијације за реакцију димеризације C протеина на (А) динамику система EcoRV, (В) брзину преласка експресије рестрикционе ендонуклеазе у укључено стање и (С) стабилност стационарног стања. Слика адаптирана из (Rodic, Vladojevic, Zdobnov, et al., 2017).

4.3.2 Смањивање ефикасности транслације c транскрипта

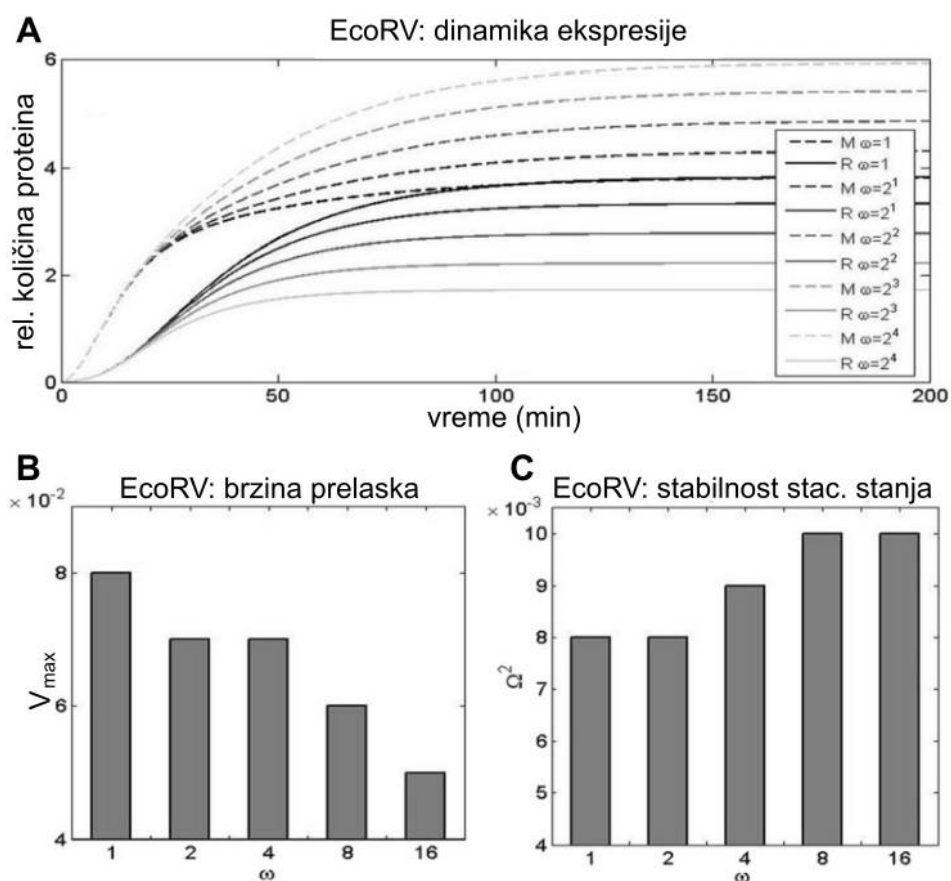
Транскрипти сва три гена система EcoRV поседују место за везивање рибозома, па су им додељене исте wt вредности константи транслације, 3 1/min . Вредност константе k_C је постепено смањивана до $3/5 \text{ 1/min}$, што одговара мањој ефикасности транслације c транскрипта у систему AhdI. Ова пертурбација има сличан, али нешто слабији ефекат на својства динамике експресије у поређењу са повећавањем $\overline{K_1}$: релативно кашњење експресије рестрикционе ендонуклеазе се не мења (Слика 17А), брзина њеног преласка из искљученог у укључено стање опадне за $\sim 50\%$ (Слика 17В), а стабилност стационарног стања система за $\sim 40\%$ (Слика 17С). Према томе, и пертурбисање k_C ремети два од три жељена динамичка својства експресије.



Слика 17. Утицај смањивања константе транслације за С протеин на (А) динамику система EcoRV, (В) брзину преласка експресије рестрикционе ендонуклеазе у укључено стање и (С) стабилност стационарног стања. Слика адаптирана из (Rodic, Blagojevic, Zdobnov, et al., 2017).

4.3.3 Повећавање афинитета за кооперативно везивање два димера С протеина

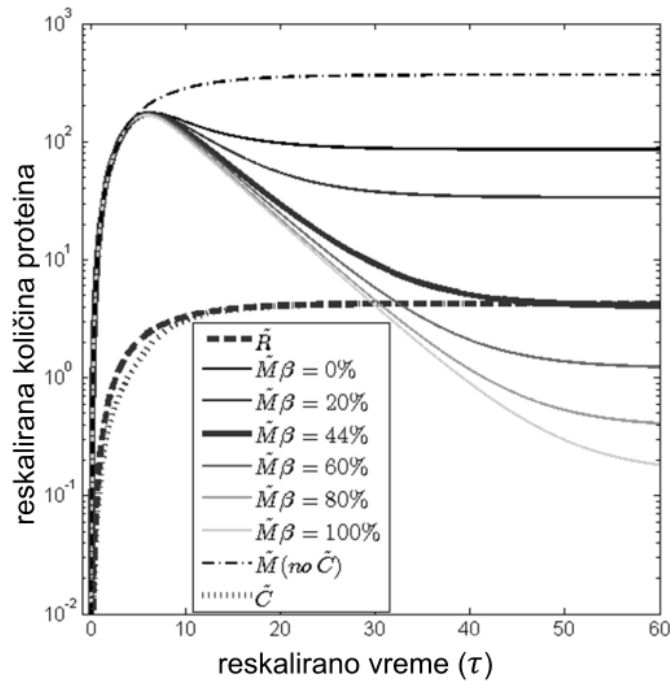
У систем EcoRV је уведена кооперативност у везивању два С димера за ДНК и постепено је повећавана (почевши од $\omega = 1$) приближавајући га систему AhdI за који је карактеристична веома висока вредност кооперативности. Ова пертурбација система, у поређењу са друге две, има нарочито упадљив ефекат на нивое метилтрансферазе и рестрикционе ендонуклеазе у стационарном стању (Слика 18А), при чему са повећањем ω расте однос њихових концентрација. Рана динамика која обухвата кашњење у експресији рестрикционе ендонуклеазе остаје иста, док v_{max} постепено опадне за $\sim 40\%$ (Слика 18В). Стабилност стационарног стања се са максималним порастом кооперативности унеколико повећа (за $\sim 25\%$, Слика 18С). Према томе, може се закључити да увођење високе кооперативности у систем EcoRV ремети једно од три жељена својства динамике експресије, док на преостала два не утиче значајно, као и да додатно има велики утицај на нивое и однос равнотежних концентрација ензима (Rodic, Blagojevic, Zdobnov, et al., 2017).



Слика 18. Утицај повећавања кооперативности у везивању С димера за ДНК на (А) динамику система EcoRV, (Б) брзину преласка експресије рестрикционе ендонуклеазе у укључено стање и (С) стабилност стационарног стања. Слика адаптирана из (Rodic, Blagojevic, Zdobnov, et al., 2017).

4.4 Предвиђање динамике експресије Р-М система Kpn2I и анализа улоге регулаторног механизма блокирања елонгације транскрипције

У (Klimuk et al., 2018) је по први пут експериментално (наши сарадници) и теоријски (наша група) окарактерисан Р-М систем регулисан и на нивоу иницијације, и на нивоу елонгације транскрипције. У моделовању и предвиђању динамике његове експресије су узети у обзир регулаторни механизми подржани резултатима експеримената, а исти резултати су искоришћени и за директну процену вредности параметара модела. Основни механизми регулације у овом систему подразумевају преклапање Р.М и Р.С због чега се РНК полимеразе не могу истовремено везати за оба промотора, као и сложену улогу С протеина везаног за ДНК низводно од Р.С који, с једне стране утишава транскрипцију m гена спречавајући везивање РНК полимеразе за Р.М, а с друге стране модулише ефикасност транскрипције s гена налазећи се на путу РНК полимерази која преписује тај ген. Теоријски предвиђена динамика успостављања wt система у ћелији (Слика 19, дебље криве) указује на брзу почетну акумулацију метилтрансферазе чији ниво након достизања максималне вредности опада до равнотежног. Такође, судећи по нагибу криве $R(t)$, може се приметити да је стационарно стање за рестрикциону ендонуклеазу достигнуто великом брзином.

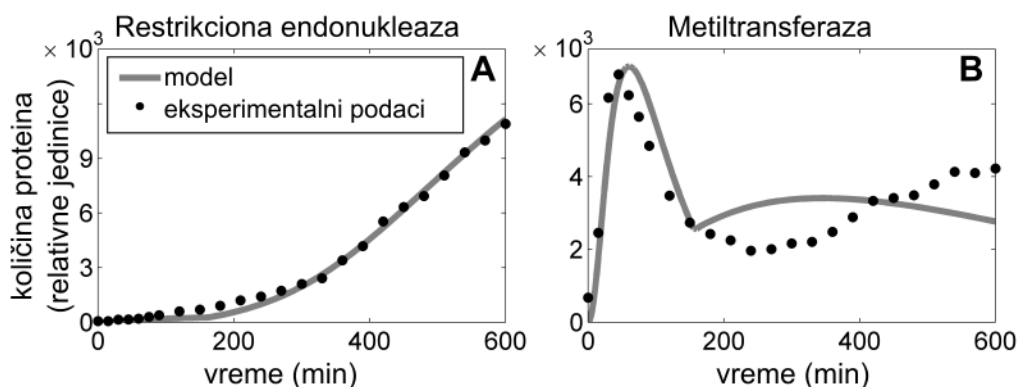


Слика 19. Предвиђено понашање Р-М система Крп2I током времена по његовом уласку у ћелију на плазмиду. Динамика wt система Крп2I коју предвиђа модел је приказана дебелим линијама – пуном је представљен ниво метилтрансферазе (М; који одговара wt вредности ефикасности елонгације, $\beta = 44\%$), цртастом ниво рестрикционе ендонуклеазе (R), а тачкастом ниво С протеина. Танком тачка-цртастом линијом је представљена динамика М у одсуству С протеина, а танким пуним линијама динамика М при различитим вредностима ефикасности елонгације која расте у смеру од тамнијих ка светлијим кривама. Ниво протеина (на у-оси) и време (на х-оси) су рескалирани као што је објашњено у поглављу Метода 3.4.3, чиме нису нарушени релативни односи протеина. Слика адаптирана из (Klimuk et al., 2018).

Значај регулације транскрипције С протеином је показан симулацијом динамике модела из којег је она елиминисана тако што је у изразима (3.114)-(3.115) за транскрипциону активност промотора задато $\bar{C}(\tau) = 0$. У овом случају количина метилтрансферазе достиже веома висок ниво који се одржава у стационарном стању и који би вероватно био токсичан за ћелију (Слика 19, тачка-цртаста крива). Због токсичности вишка метилтрансферазе, било би веома тешко или чак немогуће експериментално уклонити регулацију С протеином из система. Још би већи изазов био раздвојити ефекте везивања С протеина на транскрипцију m и s гена ради детаљнијег испитивања механизма регулације транскрипције на нивоу елонгације, али се ово може постићи теоријски, коришћењем биофизичког модела система (Klimuk et al., 2018). Вредност ефикасности елонгације транскрипције кроз барикаду коју стварају С протеини на ДНК је процењена на $\beta = 0.44$ (поглавље Метода 3.4.3), што указује на то да око $\sim 1/2$ РНК полимеразе успе да заврши транскрипцију s гена. Како би се испитали разлози ове, релативно мале ефикасности блокирања транскрипције, параметар β је вариран у моделу у опсегу од 0 до 1. Као што се може видети на Слици 19, потпуна непропусност барикаде за пролазак РНК полимеразе ($\beta = 0\%$) резултира динамиком експресије метилтрансферазе најприближнијом граничном случају када не делује никаква регулација С протеином ($\bar{C}(\tau) = 0$). Наиме, када је блокада елонгације транскрипције s гена максимално ефикасна, мала количина С протеина која се синтетише није довољна за контролу синтезе метилтрансферазе репресијом транскрипције са Р.М (Klimuk et al., 2018). Са друге стране, у недостатку ефекта блокаде који смањује експресију s гена, ниво метилтрансферазе би био превише низак, односно вероватно недовољан за постизање балансиране ензимске активности метилтрансферазе и рестрикционе ендонуклеазе у ћелији.

4.5 Предвиђање динамике експресије Р-М система Esp1396I и анализа најбољег фита експерименталних података

Наши сарадници из лабораторије проф. Северинова су мерили ниво рестрикционе ендонуклеазе и метилтрансферазе у времену по уносу Р-М система Esp1396I у наивну бактеријску ћелију и добили су динамичке обрасце експресије чији је репрезентативни пример представљен на Слици 20 (Morozova et al., 2016). Модел регулације система Esp1396I смо конструисали са циљем да испитамо да ли су постојеће знање о механизмима који делују у овом систему и минималистичка претпоставка о константним вредностима параметара довољни да се ови подаци репродукују и објасне (Morozova et al., 2016). Конкретно, узети су у обзир механизми регулације транскрипције са оба промотора С протеином (Bogdanova et al., 2009), као и механизам спрегнуте транслације цистрона гена у оквиру истог *cr* транскрипта (Cesnaviciene et al., 2003), а вредности параметара су вариране у биолошки релевантним опсезима (Phillips et al., 2012; Sneppen & Zocchi, 2005). Иако модел има релативно велики број слободних параметара (конкретно, 12) чије је вредности неопходно проценити фитовањем, његова флексибилност је ограничена независношћу експресије рестрикционе ендонуклеазе од експресије метилтрансферазе, што је искоришћено да се задатак фитовања спроведе у два корака, на одвојеним групама параметара (видети поглавље Метода 3.5.3). Поред тога, захваљујући идентичном обрасцу регулације Р.СR који је описан у системима Esp1396I и AhdI, у моделовању Esp1396I је могла да се искористи предност постојања већ провереног термодинамичког модела регулације овог промотора у систему AhdI (Bogdanova et al., 2008).

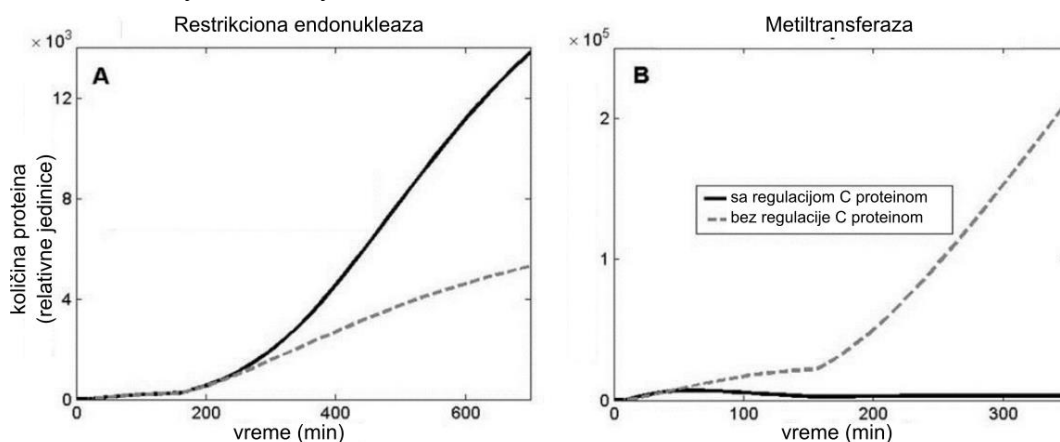


Слика 20. Предвиђање динамике експресије ензима Р-М система Esp1396I које најбоље фитује експерименталне податке. Измерене количине (у релативним јединицама) (А) рестрикционе ендонуклеазе и (Б) метилтрансферазе у времену су представљене кружићима, док је најбољи фит модела представљен пуним линијама. Слика адаптирана из (Morozova et al., 2016).

Као што се може видети на Слици 20, добијен је веома добар фит података за $R(t)$, као и података за $M(t)$ у почетном интервалу експеримента (првих ~ 160 минута). Показало се да репродуковање измерене динамике захтева високу стабилност транскрипта и протеина, односно да се њихова елиминација из система претежно одвија разблаживањем услед деобе ћелија. Овај захтев се може наслутити из приближно квадратне зависности мерене концентрације рестрикционе ендонуклеазе од времена, а накнадно је и подржан експериментом (Morozova et al., 2016). Значајно је приметити да модел може да објасни предложене основне квалитативне особине динамике успостављања Р-М система у новом домаћину: високу почетну акумулацију (пик) метилтрансферазе и кашњење почетка експресије рестрикционе ендонуклеазе. Након почетног пика, ниво метилтрансферазе је растао до краја експеримента. Модел, са друге стране, предвиђа неизбежно опадање концентрације метилтрансферазе са порастом концентрације С протеина који утишава транскрипцију са Р.М. За ово неслагање је највероватније одговорна промена неких физиолошких параметара ћелија која се наслућује из видљивог преласка са веће на мању брзину раста културе (видети Сliku 8) (Klumpp et al., 2009; Klumpp & Hwa, 2014). Промена вредности параметара са временом није

укључена у модел јер није било директних података о облику ове зависности. Нарочито утицајан би могао бити ефекат промене броја копија плаزمида који носи Р-М систем у ћелији (Graovac et al., 2019). Због некоординисаности репликације плазмида са деобом ћелија, у периоду брзе деобе ћелија (која је приближно једнака деоби плазмида) се може очекивати релативно константан број плазмида по ћелији, али када брзина деобе ћелија опадне, и даље брза репликација плазмида доводи до повећања броја копија по ћелији. Наравно, постоји и могућност да из модела одсуствује неки значајан механизам регулације који узрокује неслагање и који тек треба открити, али пре доношења таквог закључка треба испитати ефекте глобалног физиолошког стања ћелија, односно његових промена на динамику експресију Р-М система (Morozova et al., 2016).

За разлику од сложене динамике $M(t)$ која се не може у потпуности објаснити минималним моделом регулације система, измерена динамика $R(t)$ прати наизглед једноставну квадратну зависност, која делује неочекивано имајући у виду сложени механизам регулације Р.СР С протеином. Да ова динамика заиста настаје као резултат моделованог сложеног механизма регулације показује симулација динамике система у одсуству С протеина (Слика 21А), тј. када су Р.СР и Р.М само базално транскрипционо активни док је све остало у моделу, укључујући и вредности параметара одређене фитовањем, непромењено. Овакву анализу је немогуће извести експериментално због показане токсичности метилтрансферазе када је присутна у вишку (Cesnaviciene et al., 2003), што модел управо и предвиђа у одсуству репресије промотора Р.М С протеином (Слика 21В). Рестрикциона ендонуклеаза, са друге стране, без активације иначе слабог Р.СР промотора С протеином достиже знатно ниже концентрације у ћелији (Слика 21А). Чак и када би ћелије без регулације С протеином преживеле, веома висок ниво метилтрансферазе у односу на рестрикциону ендонуклеазу би резултирао брзим метиловањем стране ДНК, односно не остављањем могућности да она буде препозната од стране рестрикционе ендонуклеазе и уништена.



Слика 21. Предвиђена динамика експресије ензима Р-М система Esp1396I у одсуству С протеина. Динамика промене нивоа (А) рестрикционе ендонуклеазе и (В) метилтрансферазе је приказана у присуству (пуна линија) и у одсуству (цртаста линија) регулације транскрипције С протеином. Слика адаптирана из (Morozova et al., 2016).

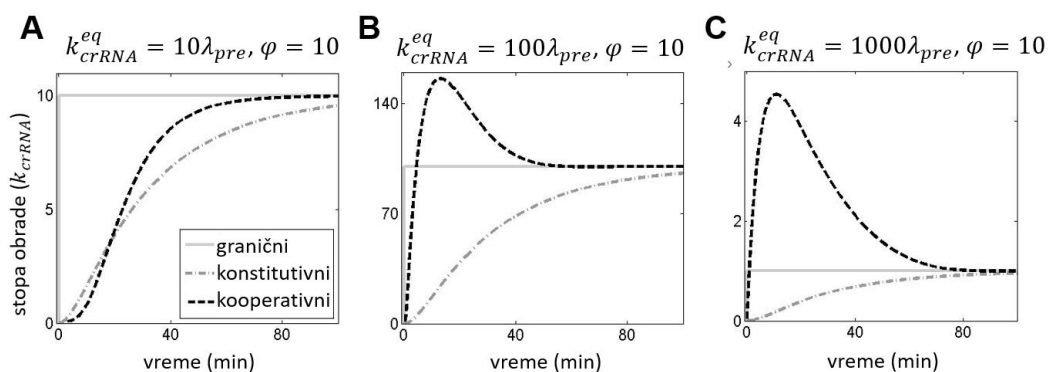
4.6 Ефекти специфичних конструктивних карактеристика система CRISPR-Cas на својства динамике његове експресије

Високо кооперативно везивање репресора за ДНК у региону промотора, конкуренција протеина активатора са протеинима репресорима за ДНК и брза неспецифична разградња pre-crRNK се истичу као специфичне конструктивне карактеристике система CRISPR-Cas. Међутим, како детаљи механизма регулације транскрипције нису довољно познати да се он може моделовати, конструисан је модел регулације система у коме је он апроксимиран добро проученим механизмом регулације Р-М система, који са њим дели сличне регулаторне принципе (Rodic,

Vlagojevic, Djordjevic, et al., 2017). Симулирањем понашања таквог система, предвиђена су основна квалитативна својства динамике експресије оригиналног система CRISPR-Cas Типа I-E. Постављањем неколико поједностављених верзија модела, циљано је анализирана улога појединих изолованих регулаторних својстава.

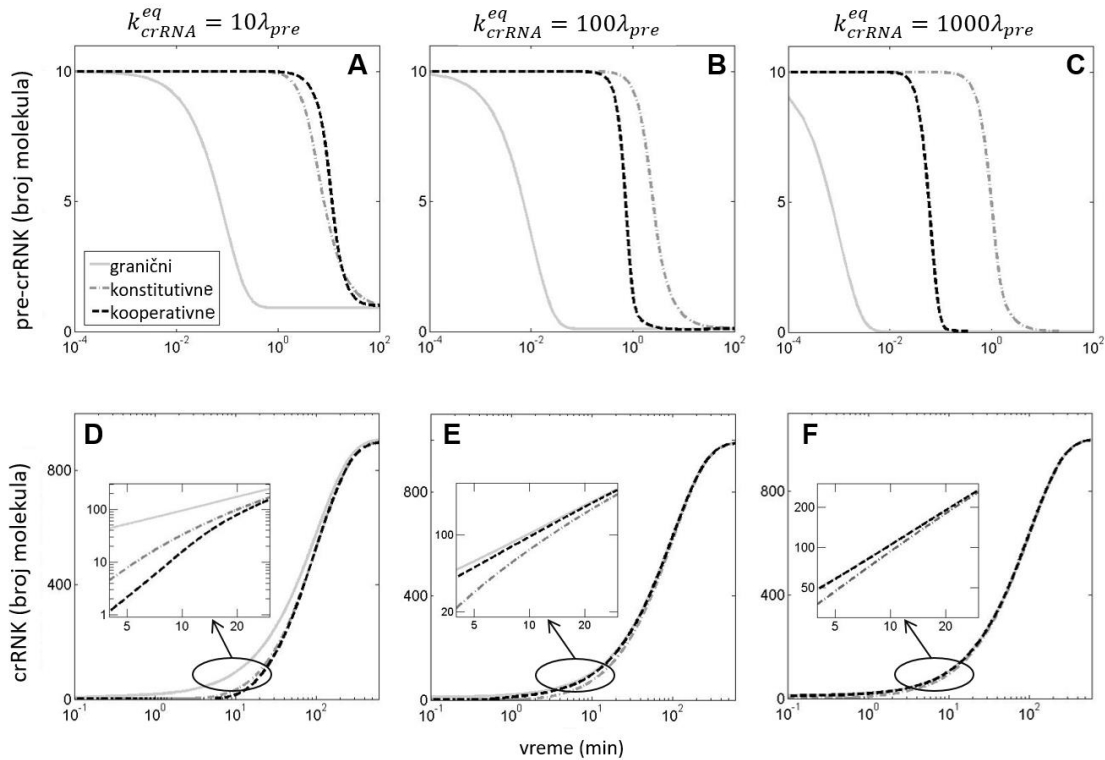
4.6.1 Ефекти постепене синтезе протеина Casбе на динамику експресије pre-crRNK и crRNK

„Конститутивним“ и „кооперативним моделом“ је уведен ефекат постепене, биолошки реалистичне синтезе протеина Casбе на динамику експресије pre-crRNK и crRNK, који је раније одсуствовао из моделованог система (Djordjevic et al., 2012), тј. била је моделована тренутна експресија огромне количине ових протеина која доводи до тренутног скока вредности стопе обраде pre-crRNK (k_{crRNA}). У овим моделима је стопа обраде пропорционална концентрацији протеина Casбе чија је експресија постепена. Истовремено је симулирана и динамика граничног модела који претпоставља тренутни скок вредности k_{crRNA} описан степенастом функцијом времена, како би продукти реалистичних модела могли да се упореде са теоријски максималном брзином индуковања обраде pre-crRNK (Слика 22). Ради изоловања ефекта обраде pre-crRNK на динамику система, претпостављена је константна транскрипциона активност P.CRISPR (10 1/min) у сва три модела.



Слика 22. Динамика промене стопе обраде pre-crRNK по покретању система при константној транскрипционој активности P.CRISPR. Криве су добијене симулацијом граничног (пуна), конститутивног (тачка-цртаста) и кооперативног модела (цртаста линија) за (А-С) три задате равнотежне вредности стопе обраде pre-crRNK. Слика адаптирана из (Rodic, Vlagojevic, Djordjevic, et al., 2017).

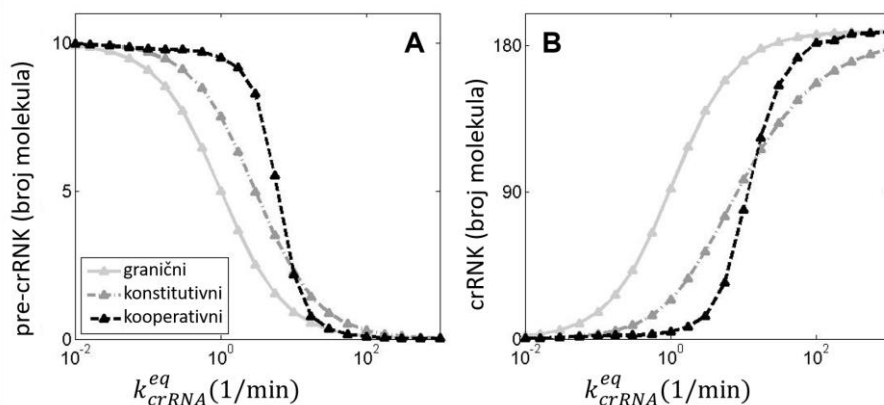
На Слици 22 видимо како се k_{crRNA} мења са временом у моделима у којима су параметри подешени тако да се у равнотежном стању оствари задата вредност k_{crRNA} у јединицама λ_{pre} (које износи 1 1/min). Приказана је динамика за три таква k_{crRNA}^{eq} ($10\lambda_{pre}$, $100\lambda_{pre}$ и $1000\lambda_{pre}$). При све три вредности k_{crRNA}^{eq} , „конститутивни модел“ продукује постепен, монотон пораст стопе обраде од почетног до равнотежног нивоа. У случају „кооперативног модела“ видимо брз почетни раст k_{crRNA} са временом праћен такође брзим повратком на равнотежну вредност при вишим задатим k_{crRNA}^{eq} ($100\lambda_{pre}$ и $1000\lambda_{pre}$), односно споријим растом према равнотежној вредности за $k_{crRNA}^{eq} = 10\lambda_{pre}$.



Слика 23. Динамика експресије pre-crRNA и crRNA по покретању система при константној транскрипционој активности P.CRISPR. За три задате равнотежне вредности стопе обраде pre-crRNA су генерисане криве експресије (A-C) pre-crRNA и (D-F) crRNA симулацијом граничног (пуна), конститутивног (тачка-цртаста) и кооперативног модела (цртаста линија). Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

Динамика промене нивоа pre-crRNA и crRNA, која одговара случајевима са Сlike 22, је приказана на Слици 23. Може се приметити да „кооперативни модел“ обезбеђује најстрмији прелаз pre-crRNA из стања велике у стање мале количине (исто али у обрнутом смеру важи за crRNA). Ако се пореде особине ране динамике при различитим задатим вредностима k_{crRNA}^{eq} , уочавају се разлике у режимима при ниским ($< 100\lambda_{pre}$) и високим вредностима ($\geq 100\lambda_{pre}$). Наиме, при ниској активности Casбе (лева колона Слике 23), crRNA се у реалистичним моделима видно спорије акумулира него у „граничном“, док при високој активности Casбе (средња и десна колона Слике 23) „кооперативни модел“ омогућава бржу акумулацију него „конститутивни“ и приближава се ефикасности „граничног модела“. Разлог постојања ових различитих режима је у бржем раном расту k_{crRNA} који „конститутивни модел“ обезбеђује при ниским, а „кооперативни“ при високим вредностима k_{crRNA}^{eq} (Слика 22).

Стационарно стање је у моделима достигнуто након >100 min (Слика 23), па та количина crRNA може бити од малог значаја за ћелију којој је ефикасна заштита од вируса потребна раније. Зато смо у нашим моделима са константном транскрипционом активношћу P.CRISPR варирали равнотежну вредност константе обраде, k_{crRNA}^{eq} (као горе), за сваку вредност симулирали динамику система и очитали количине pre-crRNA и crRNA у $t = 20$ min (Слика 24). Ове концентрације сматрамо релевантнијим за учинковиту заштиту, будући да до лизе бактеријске ћелије долази ~ 20 минута по инфицирању вирусом (Krüger & Schroeder, 1981).



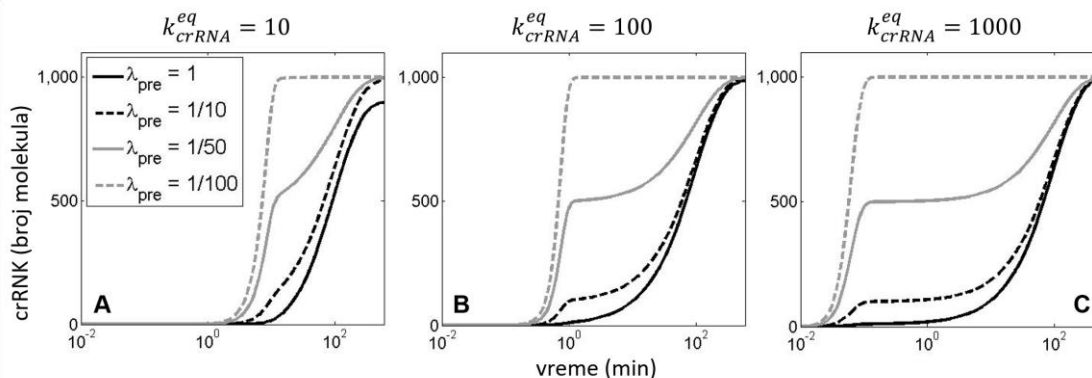
Слика 24. Ниво експресије (А) пре-сrRНК и (Б) сrRНК 20 минута по покретању система при константној транскрипционој активности P.CRISPR. Равнотежне вредности стопе обраде пре-сrRНК су вариране у граничном (пуна), конститутивном (тачка-цртаста) и кооперативном моделу (цртаста линија), док је транскрипциона активност P.CRISPR фиксирана на 10 1/min. Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

На Слици 24 се могу уочити следеће особине понашања система (Rodic, Blagojevic, Djordjevic, et al., 2017):

- 1) Понашање „кооперативног модела“ налик прекидачу, за разлику од друга два модела у којима се нивои РНК постепеније мењају са порастом k_{crRNA}^{eq} . Овакво понашање система као последицу има да се при ниским концентрацијама протеина Casбе, које могу настати услед транскрипционог „цурења“ промотора, експресија сrRНК држи у искљученом стању, тј. сrRНК је присутна у веома малој количини. Када концентрација Casбе у ћелији пређе одређену пражну вредност, систем постаје активан и брзо се произведе оптимална количина сrRНК за заштиту ћелије.
- 2) Укрштање кривих генерисаних „кооперативним“ и „конститутивним моделом“ наглашава разлике у понашању ових модела при ниским и високим стопама обраде. При ниским стопама обраде (лево од тачке укрштања), „конститутивни модел“ је ефикаснији у експресији сrRНК, тј. рано се достижу више концентрације него у „кооперативном моделу“. При високим стопама обраде (десно од тачке укрштања) које су достигнуте у експериментима (Djordjevic et al., 2012; Rougach et al., 2010) важи супротно, при чему се ниво сrRНК достигнут у „кооперативном моделу“ приближава теоријском максимуму задатом граничним моделом, који описује бесконачно брзо (тренутно) покретање обраде пре-сrRНК. Према томе, будући да се *wt* систем CRISPR-Cas одликује високо кооперативном регулацијом транскрипције, њена улога може бити у брзом обезбеђивању високе експресије сrRНК при високим стопама обраде пре-сrRНК.
- 3) Количина сrRНК довољна за мерљиву заштиту ћелије од вируса (>10 молекула, (Rougach et al., 2010)) се у „кооперативном моделу“ може добити већ при релативно ниској вредности k_{crRNA}^{eq} (~11 1/min). Због оштрог прелаза из искљученог у укључено стање, мало додатно повећање стопе обраде води великом порасту у количини сrRНК.
- 4) При високим вредностима k_{crRNA}^{eq} неизбежно долази до сатурације броја синтетисаних сrRНК, због ограничености количине доступног супстрата (пре-сrRНК), на шта нарочито јасно указује гранични модел. Дакле, количина сrRНК достиже максималну вредност која се не може превазићи даљим додавањем протеина Casбе, односно повећавањем стопе обраде пре-сrRНК.

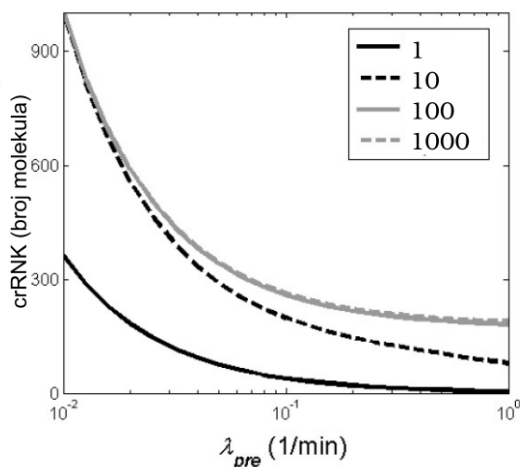
4.6.2 Ефекти смањивања стопе неспецифичне деградације pre-crRNK на динамику експресије pre-crRNK и crRNK

Брза неспецифична деградација pre-crRNK је једна од истакнутих карактеристика *wt* система CRISPR-Cas, па је испитана њена улога у контроли динамике експресије pre-crRNK и crRNK. При константној транскрипционој активности P.CRISPR, вредност λ_{pre} је постепено смањивана у моделима чији су параметри подешени тако да стопа обраде у стационарном стању достигне три задате вредности (Слика 25). На тај начин, вредност λ_{pre} је варирана у смеру *ремећења* брзе неспецифичне деградације запажене у *wt* систему CRISPR-Cas.



Слика 25. Динамика експресије crRNK по покретању система у кооперативном моделу, при константној транскрипционој активности P.CRISPR и опадајућим вредностима стопе неспецифичне деградације pre-crRNK. Приказане су криве добијене за (A-C) три задате равнотежне вредности стопе обраде pre-crRNK. Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017)

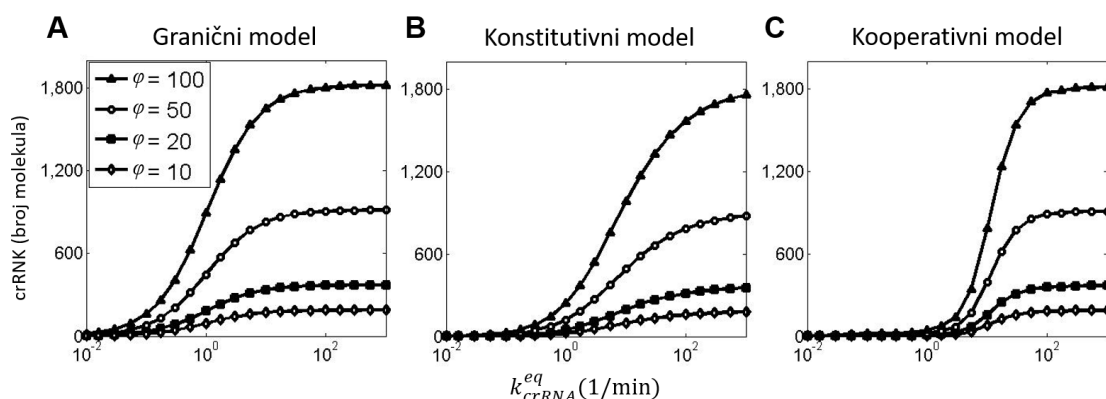
Приказани су само резултати симулације „кооперативног модела“ пошто су еквивалентне промене у динамици добијене варирањем λ_{pre} у сва три модела (Слика 25). Са успоравањем неспецифичне деградације pre-crRNK се деформише стандардни Хилов (сигмоидни) облик (Hill, 2013) криве crRNK(t) и смањује се кашњење почетка експресије crRNK за тренутком покретања експресије система (у $t = 0$ min). Услед смањеног кашњења, при нижим вредностима λ_{pre} се читава виши ниво синтетисаних crRNK у $t = 20$ min, као што показује Слика 26. Да би у $t = 20$ min била синтетисана релативно велика количина crRNK, висока вредност стопе неспецифичне деградације pre-crRNK мора бити комбинована са такође високом активношћу протеина Casбе, тј. високим k_{crRNA}^{eq} (Слика 26).



Слика 26. Ниво експресије crRNK 20 минута по покретању система при константној транскрипционој активности P.CRISPR за различите вредности стопе неспецифичне деградације pre-crRNK. Приказане су криве кооперативног модела добијене за четири задате равнотежне вредности стопе обраде pre-crRNK. Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

4.6.3 Ефекти повећавања транскрипционе активности промотора P.CRISPR на динамику експресије pre-crRNK и crRNK

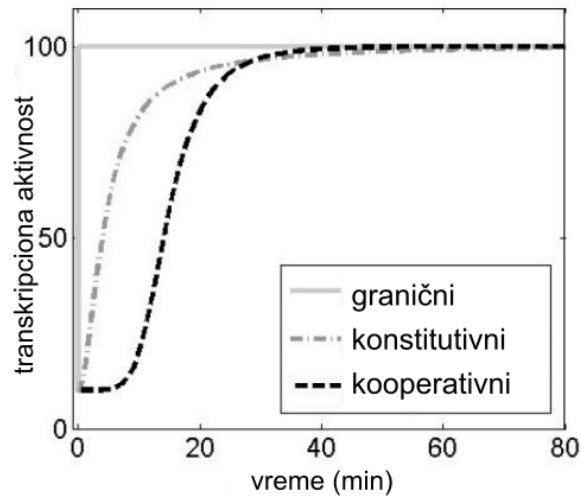
Као што је показано у поглављу 4.6.1, систем има ограничену могућност повећавања експресије crRNK путем повећавања количине протеина Casбе у ћелији, који обрађују pre-crRNK. Наиме, при одређеној, високој вредности k_{crRNA}^{eq} постиже се максимално искоришћење доступног супстрата, па се продукција crRNK може додатно повећати само ако се повећа ниво pre-crRNK. У *wt* систему CRISPR-Cas, оба промотора (и P.CCas, и P.CRISPR) су утишана кооперативно везаним протеинима H-NS за ДНК, па је разумно очекивати да сигнал за покретање експресије система, који изазива уклањање H-NS протеина са ДНК, доведе до повећања активности оба промотора. Зато је анализа, еквивалентна приказаној на Сlici 24 поновљена за различите, више вредности φ_{crispr} . Конкретно, посматрано је како концентрација crRNK у $t = 20 \text{ min}$ зависи од равнотежне стопе обраде (k_{crRNA}^{eq}) у сва три модела, за 4 различите вредности константне транскрипционе активности P.CRISPR (Слика 27).



Слика 27. Зависност нивоа експресије crRNK 20 минута по покретању система од равнотежне вредности стопе обраде pre-crRNK при растућим вредностима транскрипционе активности P.CRISPR (φ) у (A) граничном, (B) конститутивном и (C) кооперативном моделу. Напомена: Криве за $\varphi = 10$ су идентичне кривама са Сликe 24B. Слика адаптирана из (Rodic, Vladojevic, Djordjevic, et al., 2017).

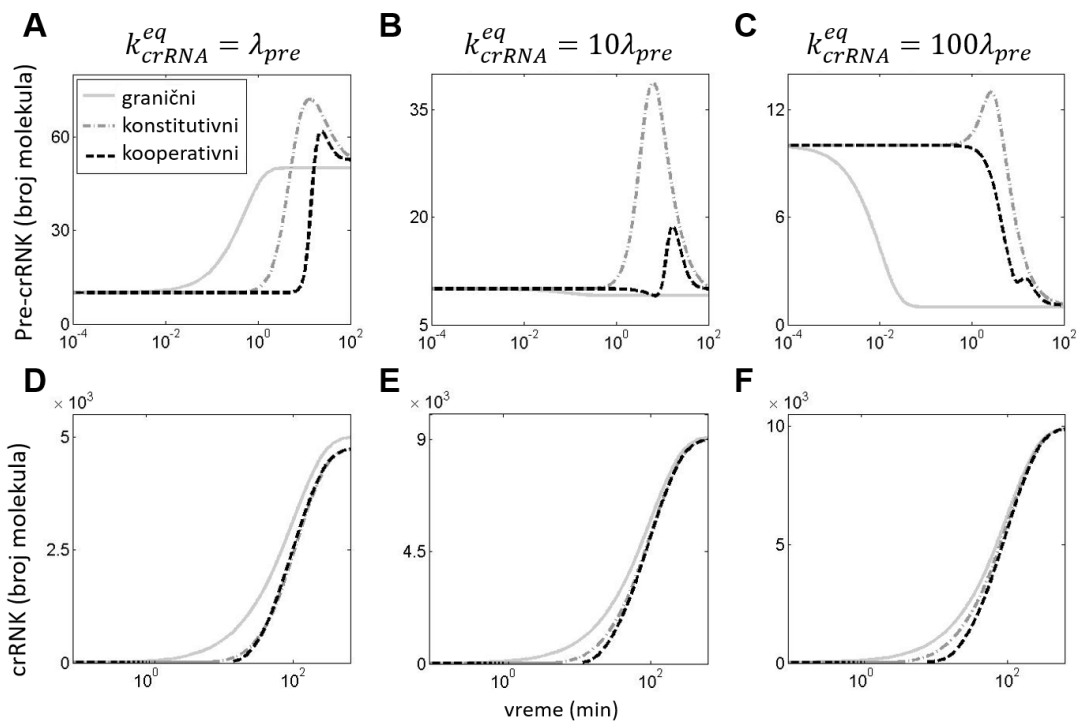
Као и на Сlici 24B, „кооперативни модел“ карактерише понашање налик прекидачу, односно најстрмији прелаз из искљученог у укључено стање експресије crRNK (Слика 27). Повећање φ_{crispr} робусно уклања сатурацију у сва три модела и омогућава експресију већих количина crRNK. Притом, исти ефекат удвостручења количине crRNK до којег долази када се k_{crRNA}^{eq} повећа за ред величине (нпр. са 10 на 100 1/min при $\varphi_{crispr} = 10 \text{ 1/min}$ у „кооперативном моделу“) се може постићи ако се φ_{crispr} повећа само 2 пута (са 10 1/min на 20 1/min при $k_{crRNA}^{eq} = 10 \text{ 1/min}$).

Коначно, испитано је и какве додатне ефекте на динамику експресије система може имати постепено покретање активности P.CRISPR која одликује стварне биолошке системе (Rodic, Vladojevic, Djordjevic, et al., 2017). Како би φ_{crispr} , по појави сигнала за покретање експресије система (у $t = 0$), постепено достигло максималну вредност карактеристичну за активну експресију система, уведена је регулација промотора P.CRISPR. Конкретно, „конститутивном моделу“ је додат једноставан механизам активације транскрипције путем везивања једног, активаторског мономера С протеина за ДНК, док је у „кооперативном моделу“ исти механизам регулације С протеином, уочен у P-M систему AhdI, примењен и на P.CCas, и на P.CRISPR. У оба случаја је, приликом подешавања вредности параметара, задато да φ_{crispr} порасте са почетне вредности 10 1/min, на крајњу 100 1/min (Слика 28). Исти услов је примењен на сваки од модела за по три различите фиксираних вредности k_{crRNA}^{eq} .



Слика 28. Динамика промене транскрипционе активности P.CRISPR у граничном (пуна), конститутивном (тачка-цртаста) и кооперативном моделу (цртаста линија). Параметри модела су подешени тако да динамика пораста φ_{crispr} не зависи од k_{crRNA}^{eq} . Слика адаптирана из (Rodic, Vladojevic, Djordjevic, et al., 2017).

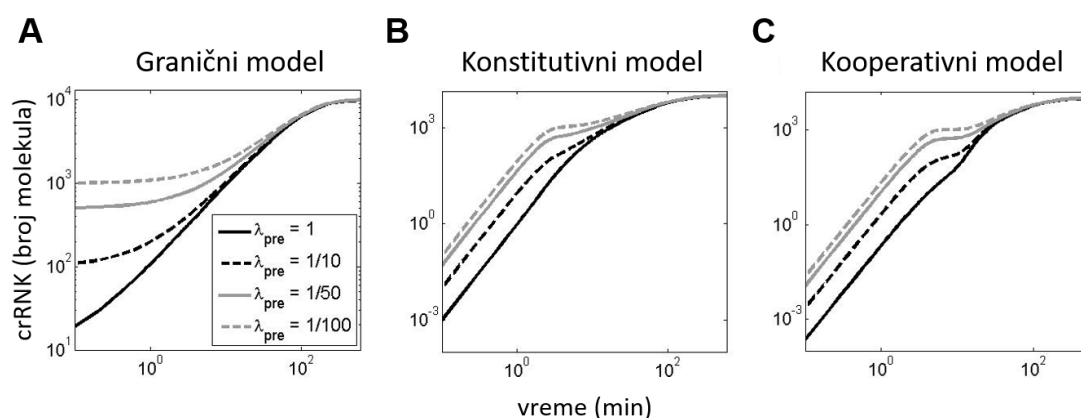
На Слици 29 је представљена добијена динамика експресије pre-crRNK и crRNK. Примећује се да „кооперативни модел“, који се ослања на комбинацију позитивне и негативне регулације транскрипције са елементом кооперативности, омогућава стабилније понашање система у поређењу са „граничним“ и „конститутивним“. Ово се, са једне стране, одражава у мањем отклону од почетне вредности који ниво pre-crRNK прави пре него што се коначно задржи на крајњој, равнотежној вредности. Са друге стране, у „кооперативном моделу“, је при различитим нивоима активности протеина Casбе, тј. различитим задатим k_{crRNA}^{eq} , најробусније очувано кашњење експресије crRNK за почетком покретања система (Слика 29).



Слика 29. Динамика експресије pre-crRNK и crRNK по покретању система у коме транскрипциона активност P.CRISPR постепено достиже 100 1/min. За три задате равнотежне вредности стопе обраде pre-crRNK су генерисане криве експресије (A-C) pre-crRNK и (D-F) crRNK симулацијом

граничног (пуна), конститутивног (тачка-цртаста) и кооперативног модела (цртаста линија). Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

Слика 29 указује на то да укључивање кооперативности у регулацију транскрипције P.CCAs и P.CRISPR доприноси одлагању почетка експресије crRNA по примању одговарајућег сигнала за активацију система, а показали смо да исти ефекат има и брза неспецифична деградација pre-crRNA (поглавље 4.6.2). Када се фиксира $k_{crRNA}^{eq} = 100 \text{ nM/min}$ и симулира динамика сва три модела (са регулацијом оба промотора) за опадајуће вредности стопе неспецифичне деградације pre-crRNA (λ_{pre}), може се видети како ове две конструктивне карактеристике заједно повећавају кашњење (Слика 30). У одсуству било какве регулације P.CRISPR („гранични модел“), количина crRNA у раном периоду брзо расте са смањивањем λ_{pre} , док „кооперативни модел“ успева да обезбеди највеће кашњење, тј. најниже нивое crRNA, у поређењу са другим моделима при свим вредностима λ_{pre} . Кашњење почетка експресије crRNA приликом активације система CRISPR-Cas до сада није постулирано као захтев значајан за функцију система, односно принцип његовог дизајна. Међутим, улога чак две структурне карактеристике система у стварању кашњења указује на могућу корист ћелије од ове особине динамике експресије crRNA (Rodic, Blagojevic, Djordjevic, et al., 2017).



Слика 30. Динамика експресије crRNA по покретању система у коме транскрипциона активност P.CRISPR постепено достиже 100 1/min , за опадајуће вредности стопе неспецифичне деградације pre-crRNA. Гранични, конститутивни и кооперативни модел су симулирани при фиксираној вредности $k_{crRNA}^{eq} = 100 \text{ nM/min}$. Слика адаптирана из (Rodic, Blagojevic, Djordjevic, et al., 2017).

5 Дискусија

Може се рећи да рестрикционо-модификациони и CRISPR-Cas системи приступају задатку одбране ћелије од стране ДНК из веома различитих углова, а и међу самим Р-М системима постоји широк спектар варијација у архитектури и механизму регулације експресије гена. Ипак, заједничка имунска функција коју обављају за ћелију нужно намеће специфичне захтеве у понашању система које они морају да задовоље одговарајућом конструкцијом. Предложили смо да се у случају Р-М система Типа II ти захтеви огледају у особинама динамике експресије ензима приликом успостављања система у ћелији као његовом новом домаћину. Конкретно, три предложена динамичка принципа дизајна Р-М система су: 1) кашњење почетка експресије рестрикционе ендонуклеазе, односно иницијално доминантна експресија метилтрансферазе, 2) брз, сигмоидни пораст нивоа рестрикционе ендонуклеазе након почетка њене синтезе и 3) стационарно стање без осцилација у концентрацијама рестрикционе ендонуклеазе и метилтрансферазе. Раније претпоставке су, поред кашњења у експресији рестрикционе ендонуклеазе, укључивале и балансиран однос ензима у стационарном стању који обезбеђује да ниједан од њих не буде токсичан за домаћина и да се одбрана од унетих ДНК може ефикасно обављати. Иако конструктивно сасвим другачији, због своје сличне имунске функције би системи CRISPR-Cas такође могли поштовати неке од ових принципа (Djordjevic, 2013; Rodic, Blagojevic, Zdobnov, et al., 2017).

Дефинисали смо квантитативне мере три предложена динамичка принципа како бисмо могли да одредимо да ли дизајн одређеног Р-М система задовољава ове принципе приликом његовог успостављања у ћелији и на који начин истакнуте конструктивне карактеристике сваког система томе доприносе (Rodic, Blagojevic, Zdobnov, et al., 2017). Р-М системи AhdI и EcoRV имају супротне карактеристике архитектуре – конвергентну, односно дивергентну – и регулације, које се односе на константе транслације и димеризације за С протеин, као и јачину кооперативне интеракције два С димера на ДНК. Међутим, предвиђена *wt* динамика успостављања оба система задовољава све динамичке принципе, што подржава нашу хипотезу и указује на то да су комбинације регулаторних својстава које се у њима срећу резултат оптимизације исте функције (Rodic, Blagojevic, Zdobnov, et al., 2017). У систему Kpn2I су преклопљени промотори гена за метилтрансферазу и С протеин, а среће се и двострука улога С протеина – у репресији иницијације транскрипције гена за метилтрансферазу и блокирању елонгације транскрипције сопственог гена. Осим што наведени механизам регулације транскрипције на нивоу елонгације није испитиван у другим Р-М системима, крајњи резултат необичне регулације система Kpn2I, тј. образац експресије гена, је тешко наслутити и због испреплетености ефеката везивања РНК полимеразе за промоторе Р.М и Р.С и С протеина за своја везивна места. Динамика експресије система Kpn2I коју предвиђа модел представља интересантан резултат јер показује да се и овом необичном регулаторном стратегијом може обезбедити велика почетна акумулација метилтрансферазе у односу на рестрикциону ендонуклеазу тј, у складу са нашом дефиницијом величине τ_{delay} , релативно кашњење неопходно за иницијалну заштиту генома бактерије-домаћина. Осим тога, уочава се и брз пораст нивоа рестрикционе ендонуклеазе до равнотежне вредности (Klimuk et al., 2018). Пик у раној експресији метилтрансферазе је уочен и приликом првог експерименталног мерења експресије ензима Р-М система *in vivo* на нивоу појединачних ћелија, за систем Esp1396I (Morozova et al., 2016). Модел регулације овог система је показао да се запажено кашњење експресије рестрикционе ендонуклеазе за почетном акумулацијом метилтрансферазе, као и пратећи брз пораст нивоа рестрикционе ендонуклеазе могу постићи описаним дизајном конструкције система (Morozova et al., 2016).

Добијена зависност мере својства динамике од параметра који контролише одређену конструктивну карактеристику система може да укаже на улогу дате карактеристике у постизању тог динамичког својства. Ако се неутралисањем кључних конструктивних карактеристика система ремете својства његове динамике, онда се та својства могу посматрати као принципи дизајна система који су наметнули присуство уочене комбинације карактеристика. Ова парадигма је примењена на моделе Р-М система AhdI и EcoRV и показано је да се постепеним укидањем сваког од карактеристичних регулаторних својстава у оба система ремети бар један динамички принцип (Rodic, Blagojevic, Zdobnov, et al., 2017). Ефикасност блокаде пута РНК полимерази током елонгације транскрипције се истиче као битна конструктивна карактеристика система Kpn2I. Варирање пропусности блокаде у моделу овог система, коју формирају С протеини везани за ДНК везивна места унутар секвенце свог гена, је довело до мењања нивоа метилтрансферазе у стационарном стању (Klimuk et al., 2018). Високо ефикасним блокирањем елонгације транскрипције *c* гена би се онемогућила довољна синтеза С протеина за репресију експресије метилтрансферазе, па би овај ензим могао да достигне концентрацију у којој би био токсичан за ћелију, док би у одсуству блокаде ниво метилтрансферазе вероватно био превише низак да парира нивоу рестрикционе ендонуклеазе. Према томе, степен ефикасности блокаде је прецизно подешен тако да обезбеди постизање оптималног равнотежног нивоа метилтрансферазе кроз интеракцију ефеката везивања С протеина на транскрипцију *m* и *c* гена. Овај резултат нарочито истиче значај, не само избора конкретних својстава регулације која су задржана у комбинацији карактеристичној за неки Р-М систем, већ и њихових специфично подешених вредности.

У складу са тим, динамика једног система се не може учинити оптималнијом увођењем кључних регулаторних својстава другог система, јер се тиме нарушава конструкција читавог система чији су делови и карактеристике током еволуције задржани зарад њиховог заједничког и синергистичког дејства (Rodic, Blagojevic, Zdobnov, et al., 2017). У прилог томе, у систему AhdI се јасно уочавају супротстављене улоге ниских ефикасности синтезе С протеина транслацијом и његове димеризације, са једне стране, и високе кооперативности у везивању два С димера за ДНК. Док су прва својства неопходна за постизање кашњења и брзог раста у динамици експресије рестрикционе ендонуклеазе, али истовремено стварају услове за каснију сувишну акумулацију овог ензима, друго је задужено за ограничавање количине рестрикционе ендонуклеазе у стационарном стању и његову стабилност (Rodic, Blagojevic, Zdobnov, et al., 2017). Занимљиво је да је у систему Esp1396I, који користи исти механизам регулације промотора P.CR С протеином, слабија кооперативност у односу на исту у систему AhdI комбинована са нешто већим афинитетом С протеина за формирање хомодимера и ефикаснијом транслацијом *c* цистрона (Bogdanova et al., 2009; Morozova et al., 2016; Rodic, Blagojevic, Zdobnov, et al., 2017). Могуће је да је комбинација ова три регулаторна својства у свим конвергентним системима у којима се јавља оптимизована по сличном принципу, како би се у стационарном стању обезбедила одговарајућа концентрација рестрикционе ендонуклеазе (Bogdanova et al., 2008; Rodic, Blagojevic, Zdobnov, et al., 2017).

О оптималним нивоима метилтрансферазе и рестрикционе ендонуклеазе у стационарном стању, тј. о њиховом оптималном односу се и раније дискутовало у литератури као о важном функционалном захтеву приликом експресије Р-М система (Morozova et al., 2016). Познато је да при прениској концентрацији метилтрансферазе, или превисокој концентрацији рестрикционе ендонуклеазе, геномска ДНК може бити фрагментисана, што би довело до смрти ћелије (Ichige & Kobayashi, 2005). Са друге стране, при превисокој концентрацији метилтрансферазе, или прениској рестрикционе ендонуклеазе, вирусна ДНК би лако могла да избегне механизам одбране и стекне заштитне метил-групе, а сувишна метилација ДНК би могла бити и летална по домаћина (Enikeeva et al., 2010). Треба напоменути да, осим од брзина синтезе и уклањања самих ензима, оптималност њиховог односа зависи и од њихових специфичних ензимских активности, потребе за димеризацијом рестрикционе ендонуклеазе, као и њене ефикасности у сечењу хемиметиловане ДНК (M. O. Nagornykh et al., 2008), па

потребне концентрације ензима и њихов однос могу донекле варирати код различитих Р-М система (Rodic, Blagojevic, Zdobnov, et al., 2017). Сигурно је да, због наведених штетних последица, Р-М системи не могу да толеришу велике разлике у нивоима самих ензима у ћелији, па је контрола њихове експресије неопходна. Нарочито изражене промене нивоа ензима и њиховог односа, у смеру пораста концентрације метилтрансферазе и смањења концентрације рестрикционе ендонуклеазе, су добијене као резултат ремећења кључних регулаторних својстава система EcoRV, тј. увођења вредности које су карактеристичне за конвергентни систем AhdI (Rodic, Blagojevic, Zdobnov, et al., 2017). Преклопљени промотори Р.СР и Р.М ниског и високог афинитета обезбеђују робусно испољавање кашњења почетка експресије рестрикционе ендонуклеазе у овом систему, међутим, изгледа да чврста повезаност и испреплетеност различитих регулаторних процеса у дивергентном систему чини стационарно стање система осетљивијим на пертурбације. Показан значај финог подешавања јачине ефекта блокаде који у систему Крп2I контролише елонгацију транскрипције *c* гена је у складу са овом претпоставком, ако се она примени на *t* и *c* гене чија је транскрипција чврсто спрегнута у овом систему.

Ауторегулација транскрипције *c* гена која обухвата активацију и репресију, а директно је имплементирана у системима AhdI, Esp1396I и EcoRV, је на неочекиван начин присутна и у систему Крп2I (Klimuk et al., 2018). Овде С протеини везани за С-бокс унутар *c* гена инхибирају експресију свог гена спречавањем дела РНК полимеразе да доврше његову транскрипцију, а индиректно је и активирају тиме што ометају везивање РНК полимеразе за промотор Р.М који је преклопљен и у конкуренцији са промотором Р.С. Притом је пропусност блокаде фино подешена да оба ефекта буду довољно изражена. Ова запажања указују на важност постојања механизма активације и репресије промотора *c* гена, којима се регулише ниво С протеина од којег даље зависи ниво једног или оба ензима и однос који се успоставља у стационарном стању (Klimuk et al., 2018). Уочен сличан принцип регулације промотора у систему CRISPR-Cas, где су носиоци функција високо кооперативне репресије транскрипције, као и њене активације уз присуство саморегулације различити глобални регулатори, концептуално приближава Р-М системима њима несродну групу бактеријских имунских система, CRISPR-Cas. Ова сличност – конкретно, између система CRISPR-Cas Типа I-E и Р-М система AhdI – је уздигнута на ниво општег принципа регулације транскрипције у овим системима у нашој тврдњи да се квалитативне особине динамике експресије система CRISPR-Cas могу предвидети анализом понашања хибридног система у коме је непозната динамика транскрипционе активности његових промотора премошћена оном коју генерише концептуално сличан механизам регулације из Р-М система AhdI.

Као кључна регулаторна својства система CRISPR-Cas Типа I-E истичу се брза деградација pre-crRNA молекула неидентификованом рибонуклеазом и високо кооперативна репресија оба промотора коју може уклонити активатор транскрипције, а која је посебно снажна у случају промотора Р.Сас (Rodic, Blagojevic, Djordjevic, et al., 2017). Резултати симулације нашег модела система који поседује оба својства показују да присуство комплетног механизма регулације транскрипције омогућава пожељне особине понашања система оптимизованог за режим активације са релативно високим концентрацијама протеина Casбе, односно високом стопом обраде pre-crRNA у crRNA. При таквим експерименталним концентрацијама је систем био успешно активиран *in vitro* (Djordjevic et al., 2012; Pougach et al., 2010), па је могуће да су приближне вредности релевантне и за његову природну активацију. Конкретно, овако регулисан систем би се понашао слично прекидачу – при концентрацијама Casбе испод одређене празне вредности, које би могле да варирају услед транскрипционог „дурења“ промотора Р.Сас, би остао неактиван, док би по достизању празне вредности концентрације Casбе прешао у режим веома брзе производње crRNA приближавајући се теоријском максимуму ове брзине (Rodic, Blagojevic, Djordjevic, et al., 2017). Већ самом активацијом промотора Р.Сас би се довољан ниво crRNA могао достићи за релативно кратко време, а истовременом активацијом Р.CRISPR би капацитет производње био веома ефикасно повећан.

Будући да су у систему CRISPR-Cas I-E оба промотора утишана везивањем протеина H-NS, вероватно се истовремено активирају, а слабија репресија промотора P.CRISPR се може објаснити великим ефектом који мало повећање његове транскрипционе активности има на ниво генерисаних crRNK (Djordjevic et al., 2012; Pougach et al., 2010; Pul et al., 2010; Rodic, Vladojevic, Djordjevic, et al., 2017; Westra et al., 2010). Приметили смо да је почетак брзе ране производње crRNK, неочекивано, одложен у односу на тренутак почетка експресије P.CCAs и да овом кашњењу специфично доприноси механизам регулације транскрипције. Како је већ познато да овај механизам у P-M систему обезбеђује кашњење почетка експресије рестрикционе ендонуклеазе, на основу сличности регулације система CRISPR-Cas се донекле могло очекивати да и у њему постоји кашњење, али занимљиво је да ово није једини механизам који га промовише у овом систему. Наиме, показали смо да и брза неспецифична деградиција pre-crRNK уводи кашњење у динамику активације система, што још снажније подржава претпоставку да је почетно одлагање експресије crRNK потребно и корисно за функцију система. Једна улога коју би кашњење могло да има је у олакшавању процеса адаптације са прајмером (енг. *primed adaptation*) (Künne et al., 2016, p. 3; Musharova et al., 2017; Severinov et al., 2016). У овом специфичном процесу адаптације, постојећи спејсер се користи као прајмер за селектовање фрагмената вирусне ДНК који ће бити уграђени у CRISPR низ као нови спејсери. На овај начин се могу сакупити узорци мутиране ДНК вируса који би, у случају коришћења постојећих спејсера, избегли CRISPR интерференцију (Sternberg et al., 2016). Показано је да се успоравањем интерференције промовише адаптација са прајмером (Künne et al., 2016, p. 3; Musharova et al., 2017; Severinov et al., 2016), па можемо претпоставити да би одлагање експресије crRNK помогло у томе.

Иако су специфични регулаторни механизми несумњиво најодговорнији за експресију гена, односно протеина које они кодирају, не треба заборавити ни ефекте физиолошког контекста који могу да доведу до великих одступања понашања система од предвиђеног (Klumpp & Hwa, 2014; Vilar et al., 2003). Наиме, услови раста диктирају у каквом ће метаболичком стању ћелија бити, како ће распоређивати своје ресурсе за експресију гена и синтезу протеина, којом брзином ће се делити итд. Из тих разлога, параметри, као што су расположивост слободних РНК полимераза и рибозома, брзине деградиције иРНК и протеина, број копија гена лоцираног на хромозому, а нарочито када је лоциран на плазмиду, брзина разблаживања концентрације молекула услед деобе ћелије, па и запремина ћелије, могу да се мењају са временом ако током посматраног периода услови раста бактерија нису константни (Klumpp et al., 2009; Klumpp & Hwa, 2014). Због тога не треба журити са закључком да је неслагање опадања нивоа метилтрансферазе у каснијем интервалу динамике, које предвиђа модел система Esp1396I, са експериментално уоченим порастом последица непознатог регулаторног механизма који модел не описује (Morozova et al., 2016). У нашем новијем истраживању смо испитали да ли се подаци могу у потпуности репродуковати ако се у модел дода ефекат променљивог броја копија плаزمида по ћелији услед неусаглашених брзина репликације плазмида и деобе ћелија (Graovac et al., 2019). Ово смо урадили након напомене наших сарадника да је број плазмида по ћелији на самом почетку експеримента био 1, а на крају већи за ~2 реда величине. Изабрали смо да моделујемо брзину репликације плазмида по узору на регистровану динамику деобе ћелија како бисмо избегли да учинимо модел непотребно сложеним – дозволили смо да брзина репликације плазмида има једну вредност на почетку, а да затим у одређеном тренутку независном од раста културе добије другу вредност. Показали смо да је ова модификација модела неопходна и довољна да се у потпуности објасне експериментални подаци (Graovac et al., 2019). Уз то смо приказали и додатне занимљиве резултате који илуструју величину утицаја динамике популације на финалну динамику експресије протеина у ћелији, који може резултирати не само квантитативним, већ и квалитативним променама динамике коју би произвео изолован регулаторни механизам. Према томе, динамику експресије система у ћелији *in vivo* треба посматрати као резултат интеракције ефеката регулације гена и глобалних

физиолошких фактора (Berthoumieux et al., 2013; Gerosa et al., 2013; Klumpp et al., 2009; Stefan et al., 2015).

6 Закључак

Истраживањем представљеним у овој тези су расветљени неки од универзалних принципа дизајна Р-М система, који су проистекли из специфичних захтева за безбедним и ефикасним обављањем њихове функције одбране ћелије од стране ДНК. Принципе дизајна смо формулисали кроз пожељне особине динамике експресије ензима Р-М система у бактеријској ћелији, чији се ниво испољености може квантификовати мерама које смо дефинисали. Анализом понашања модела четири репрезентативна Р-М система, који се одликују различитим архитектурама и регулаторним својствима, показали смо да се у великој разноврсности и привидној насумичности конструктивних карактеристика које се срећу код различитих представника ове категорије бактеријских имунских система могу препознати правила комбиновања којима се обезбеђују динамичка својства као што су почетно кашњење и каснији брз раст експресије рестрикционе ендонуклеазе, као и одржавање балансираног односа ензима у дугорочном режиму. Наши резултати указују на то да би се слични принципи могли односити и на дизајн потпуно другачије групе система, CRISPR-Cas, који представља напредно решење за обављање исте имунске функције.

Анализом одабраних Р-М система је показано да се универзални динамички принципи могу задовољити оптималним комбиновањем конструктивних карактеристика као што су преклопљени промотори, стратешки постављена места за везивање димера С протеина које може бити модулисано његовим афинитетом за димеризацију и јачином кооперативне интеракције, смањена ефикасност транслације услед одсуства места за везивање рибозома или формирања РНК укоснице која га чини тешко доступним и др. На примеру система Kpn2I смо, заједно са нашим сарадницима, у библиотеку конструктивних елемената и карактеристика Р-М система додали везивање С протеина унутар секвенце гена, са ефектом блокирања елонгације транскрипције, а с обзиром на велику разноврсност Р-М система која се уочава већ на нивоу организације гена, може се очекивати да има још таквих елемената које тек треба открити и окарактерисати. Постављеним моделима би се могао додати и модел неког од Р-М система са линеарном организацијом гена који немају С протеин, али за регулацију могу користити антисенс РНК или механизме на нивоу транслације (Mruk et al., 2011; M. Nagornykh et al., 2011; Rodic, Blagojevic, Zdobnov, et al., 2017). Демонстрирана способност оваквих система да остваре неке или све предложене динамичке принципе би додатно учврстила нашу хипотезу. Такође, нашом анализом нису обухваћене пертурбације више регулаторних својстава одједном, а таквом, сложенијом анализом *in silico* би могао да се испита степен редувантности њихових ефеката, као и могућност да се систем са једним модификованим регулаторним својством „спасе“ модификацијом неког другог (Rodic, Blagojevic, Zdobnov, et al., 2017).

Главни ограничавајући фактори у нашој теоријској анализи понашања бактеријских имунских система су доступност експерименталних података и информација о конкретном систему и могућност, односно прилика да се предвиђања модела провере експериментално. Приликом конструисања модела смо тежили да укључимо све елементе и процесе система који су неопходни да се његова регулација и експресија у ћелији адекватно опишу, уз истовремено вођење рачуна да модел не постане толико сложен да се његова параметризација не може извести помоћу доступних података. Прављењем овог компромиса се представа биолошког система дата моделом неизбежно, у одређеној мери, удаљава од стварне, али овај недостатак се толерише на рачун предности које анализа *in silico* пружа у односу на експерименталну. Сврха моделовања, да се испитају квалитативне особине динамике експресије, тј. динамички принципи дизајна система, оправдава коришћење релативно једноставних модела, а њиховом

симулацијом и анализом се може доћи до корисних сазнања о улози појединих елемената регулације у постизању жељених особина динамике. Еквивалентна експериментална анализа би подразумевала захтевне задатке конструисања система са различитим мутацијама предвиђених и локализованих ефеката, као и регистровања динамике експресије молекула система у појединачним ћелијама. Резултати наших анализа *in silico* указују на основне правце у којима се могу дизајнирати експерименти ради њихове провере и примене. И сама провера наших предвиђања захтева сложене експерименталне поставке будући да динамика експресије молекула система треба да се прати на нивоу појединачних ћелија у што уједначенијим условима раста културе, али све моћније технике флуоресцентне микроскопије и микрофлуидике дају могућност за извођење таквих мерења са високом резолуцијом и степеном контроле (Elowitz & Leibler, 2000; Longo & Hasty, 2006; Rosenfeld et al., 2005; Young et al., 2011).

Као један од изазовнијих циљева у модерној науци се појавио приступ „обрнутог инжењерства“ којим би требало реконструисати механизме регулације на основу информација о експресији гена – у екстремном случају, само из динамике експресије протеина (Barzel et al., 2015; de Jong & Geiselman, 2015; Klumpp & Hwa, 2014; Porreca et al., 2010; Stefan et al., 2015). У том процесу, неопходно је раздвојити неизбежне глобалне физиолошке ефекте, односно ефекте популације, од регулаторних. Експериментални метод којем се углавном прибегава је истовремено праћење експресије гена са регулисаног и са конститутивног промотора и, затим, поређење и анализа добијених података. На примеру модела регулације *lac* оперона је показано да су ефекти популације неопходни за објашњавање експериментално добијених података (Vilar et al., 2003). Резултати приказани у овој тези су дали само наговештај значаја који за динамику имају ефекти физиолошког контекста у коме се одвија експресија одређеног Р-М система. Ти ефекти су делимично испитани у наставку нашег истраживања, али теоријска анализа таквих придружених ефеката средине, којима се могу додати и стохастичке варијације молекуларних процеса, би захтевала велики број додатних информација из експеримената.

Подложност динамике експресије Р-М система променама изазваним ефектима динамике популације, описана у случају система Esp1396I, отвара питање робусности предложених динамичких принципа. Како те принципе сматрамо суштински важним за опстанак ћелије домаћина, а и генетичке информације самог система приликом његовог првог успостављања, можемо претпоставити да су они *in vivo* довољно отпорни на променљивост услова средине, односно да је и робусно испољавање жељених особина динамике обезбеђено дизајном њихове регулације. Идентификовањем и анализом регулаторних својстава која су у Р-М системима одговорна за робусност њихове динамике експресије би се значајно допринело разумевању регулације експресије генских мрежа, будући да је робусност понашања биолошких система од значаја за све функције које се обављају у променљивим условима средине. На пример, она је кључна за успешност лекова, будући да они ефективно пертурбишу систем на који делују, па се разумевањем начина имплементације робусности може унапредити дизајн нових лекова (Kitano, 2007). И сами Р-М системи би могли наћи примену као средства за борбу против мулти-резистентних сојева бактерија, будући да постоје докази о повезаности губитка Р-М система и олакшаног развоја резистенције на антибиотике код патогених бактерија (V. J. Price et al., 2016). За такву примену Р-М система је од суштинског значаја претходно добро разумевање њиховог понашања описаног кључним динамичким принципима међу којима је највероватније и робусност (Klumpp & Hwa, 2014).

Треба напоменути да су у овој тези динамичка својства експресије Р-М и CRISPR-Cas система интерпретирана у контексту њихове основне, имунске функције, али описане су и друге функције које они обављају у бактеријама, па би требало проверити и како се показани динамички принципи уклапају у захтеве тих других улога. На пример, транскрипти без наменског места за везивање рибозома који, услед спорије транслације, одлажу појаву С протеина у ћелији, могу подлећи интензивнијој транслацији у специфичним физиолошким

условима (Moll et al., 2002; Vesper et al., 2011). CRISPR-Cas Типа I-E у *E. coli* је у стандардним условима утишан. Поред могућности да су за његову активацију неопходни посебни услови, предложено је и да би он у *E. coli* могао имати неку другу примарну функцију уместо одбране од вируса. Због тога би требало сагледати корист коју његове предвиђене особине динамике могу донети ћелији и из других углова (Rodic, Blagojevic, Djordjevic, et al., 2017; Touchon et al., 2011). Хибридни систем који смо моделовали како бисмо дали предвиђања основних особина експресије CRISPR-Cas приликом његове активације би могао да се конструише експериментално, стављањем CRISPR низа и *cas* гена под контролу промотора P_{CR} из система AhdI и увођењем *c* и *cas* гена у ћелију на плазмиду или вирусном вектору чиме би се покренула активација експресије система. На овај начин би могла да се провере предвиђања модела – да ли динамику експресије crRNK заиста одликују кашњење и понашање налик прекидачу – али и да се постигне веродостојна имитација природне активације CRISPR-Cas I-E у *E. coli* која се може искористити за проучавање његових других функција у ћелији.

Резултати проучавања P-M система, као природних, компактних и независно регулисаних генских мрежа, имају велики потенцијал за примену у синтетичкој биологији, у конструисању вештачких генских кола за контролисано обављање одређених функција у живом систему (Wang et al., 2016; Zhang et al., 2016). Један од основних изазова у овом пољу је креирање биолошких уређаја који се понашају предвидљиво не само *in vitro*, већ и у променљивим условима *in vivo*. У одабиру архитектуре и компоненти које ће омогућити такво понашање могу значајно да помогну сазнања о структури и функцији природних генских мрежа, оптимизованих у складу са општим принципима њиховог еволуционог дизајна (Morrison et al., 2021; Savageau, 1977; Venturelli et al., 2016). P-M системи, са широким спектром компоненти и својстава регулације које користе, а независни од регулаторне машинерије ћелије домаћина, представљају одличан модел систем за ове циљеве. Према томе, поред значаја за разумевање општих принципа који повезују структуру и функцију бактеријских имунских система, резултати анализе P-M и CRISPR-Cas система представљени у овој тези се могу искористити и у примењеним експерименталним истраживањима за конструкцију генских кола.

7 Литература

- Aguda, B., Friedman, A., & Visiting Associate Professor, T. O. S. U. (2008). *Models of Cellular Regulation*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198570912.001.0001>
- Al-Attar, S., Westra, E. R., van der Oost, J., & Brouns, S. J. J. (2011). Clustered regularly interspaced short palindromic repeats (CRISPRs): The hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biological Chemistry*, 392(4), 277–289. <https://doi.org/10.1515/BC.2011.042>
- Alon, U. (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits* (1st edition). Chapman and Hall/CRC.
- Anton, B. P., Heiter, D. F., Benner, J. S., Hess, E. J., Greenough, L., Moran, L. S., Slatko, B. E., & Brooks, J. E. (1997). Cloning and characterization of the BglII restriction-modification system reveals a possible evolutionary footprint. *Gene*, 187(1), 19–27. [https://doi.org/10.1016/s0378-1119\(96\)00638-5](https://doi.org/10.1016/s0378-1119(96)00638-5)
- Asakura, Y., & Kobayashi, I. (2009). From damaged genome to cell surface: Transcriptome changes during bacterial cell death triggered by loss of a restriction-modification gene complex. *Nucleic Acids Research*, 37(9), 3021–3031. <https://doi.org/10.1093/nar/gkp148>
- Ay, A., & Arnosti, D. N. (2011). Mathematical modeling of gene expression: A guide for the perplexed biologist. *Critical Reviews in Biochemistry and Molecular Biology*, 46(2), 137–151. <https://doi.org/10.3109/10409238.2011.556597>
- Baev, M. V., Baev, D., Radek, A. J., & Campbell, J. W. (2006). Growth of Escherichia coli MG1655 on LB medium: Determining metabolic strategy with transcriptional microarrays. *Applied Microbiology and Biotechnology*, 71(3), 323–328. <https://doi.org/10.1007/s00253-006-0392-8>
- Ball, N. J., McGeehan, J. E., Streeter, S. D., Thresh, S.-J., & Kneale, G. G. (2012). The structural basis of differential DNA sequence recognition by restriction-modification controller proteins. *Nucleic Acids Research*, 40(20), 10532–10542. <https://doi.org/10.1093/nar/gks718>
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, N.Y.)*, 315(5819), 1709–1712. <https://doi.org/10.1126/science.1138140>
- Bart, A., Dankert, J., & van der Ende, A. (1999). Operator sequences for the regulatory proteins of restriction modification systems. *Molecular Microbiology*, 31(4), 1277–1278. <https://doi.org/10.1046/j.1365-2958.1999.01253.x>
- Barzel, B., Liu, Y.-Y., & Barabási, A.-L. (2015). Constructing minimal models for complex system dynamics. *Nature Communications*, 6, 7186. <https://doi.org/10.1038/ncomms8186>
- Beletskaya, I. V., Zakharova, M. V., Shlyapnikov, M. G., Semenova, L. M., & Solonin, A. S. (2000). DNA methylation at the CfrBI site is involved in expression control in the CfrBI restriction-modification system. *Nucleic Acids Research*, 28(19), 3817–3822. <https://doi.org/10.1093/nar/28.19.3817>
- Berthoumieux, S., de Jong, H., Baptist, G., Pinel, C., Ranquet, C., Ropers, D., & Geiselmann, J. (2013). Shared control of gene expression in bacteria by transcription factors and global

physiology of the cell. *Molecular Systems Biology*, 9, 634.
<https://doi.org/10.1038/msb.2012.70>

- Bickle, T. A., & Krüger, D. H. (1993). Biology of DNA restriction. *Microbiological Reviews*, 57(2), 434–450. <https://doi.org/10.1128/mr.57.2.434-450.1993>
- Bintu, L., Buchler, N. E., Garcia, H. G., Gerland, U., Hwa, T., Kondev, J., & Phillips, R. (2005). Transcriptional regulation by the numbers: Models. *Current Opinion in Genetics & Development*, 15(2), 116–124. <https://doi.org/10.1016/j.gde.2005.02.007>
- Bogdanova, E., Djordjevic, M., Papapanagiotou, I., Heyduk, T., Kneale, G., & Severinov, K. (2008). Transcription regulation of the type II restriction-modification system AhdI. *Nucleic Acids Research*, 36(5), 1429–1442. <https://doi.org/10.1093/nar/gkm1116>
- Bogdanova, E., Zakharova, M., Streeter, S., Taylor, J., Heyduk, T., Kneale, G., & Severinov, K. (2009). Transcription regulation of restriction-modification system Esp1396I. *Nucleic Acids Research*, 37(10), 3354–3366. <https://doi.org/10.1093/nar/gkp210>
- Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading, England)*, 151(Pt 8), 2551–2561. <https://doi.org/10.1099/mic.0.28048-0>
- Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., & Rimsky, S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nature Structural & Molecular Biology*, 14(5), 441–448. <https://doi.org/10.1038/nsmb1233>
- Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., & van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science (New York, N.Y.)*, 321(5891), 960–964. <https://doi.org/10.1126/science.1159689>
- Buchler, N. E., Gerland, U., & Hwa, T. (2003). On schemes of combinatorial transcription logic. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9), 5136–5141. <https://doi.org/10.1073/pnas.0930314100>
- Butler, D., & Fitzgerald, G. F. (2001). Transcriptional analysis and regulation of expression of the ScrFI restriction-modification system of *Lactococcus lactis* subsp. *Cremoris* UC503. *Journal of Bacteriology*, 183(15), 4668–4673. <https://doi.org/10.1128/JB.183.15.4668-4673.2001>
- Casadesús, J., & Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiology and Molecular Biology Reviews: MMBR*, 70(3), 830–856. <https://doi.org/10.1128/MMBR.00016-06>
- Cesnaviciene, E., Mitkaite, G., Stankevicius, K., Janulaitis, A., & Lubys, A. (2003). Esp1396I restriction-modification system: Structural organization and mode of regulation. *Nucleic Acids Research*, 31(2), 743–749. <https://doi.org/10.1093/nar/gkg135>
- Chen, C. C., Fang, M., Majumder, A., & Wu, H. Y. (2001). A 72-base pair AT-rich DNA sequence element functions as a bacterial gene silencer. *The Journal of Biological Chemistry*, 276(12), 9478–9485. <https://doi.org/10.1074/jbc.M010501200>
- Christensen, L. L., & Josephsen, J. (2004). The methyltransferase from the LlaDII restriction-modification system influences the level of expression of its own gene. *Journal of Bacteriology*, 186(2), 287–295. <https://doi.org/10.1128/JB.186.2.287-295.2004>
- de Jong, H., & Geiselman, J. (2015). Fluorescent Reporter Genes and the Analysis of Bacterial Regulatory Networks. In O. Maler, Á. Halász, T. Dang, & C. Piazza (Eds.), *Hybrid Systems Biology* (pp. 27–50). Springer International Publishing. https://doi.org/10.1007/978-3-319-27656-4_2

- Díez-Villaseñor, C., Almendros, C., García-Martínez, J., & Mojica, F. J. M. (2010). Diversity of CRISPR loci in *Escherichia coli*. *Microbiology (Reading, England)*, *156*(Pt 5), 1351–1361. <https://doi.org/10.1099/mic.0.036046-0>
- Djordjevic, M. (2013). Modeling bacterial immune systems: Strategies for expression of toxic - but useful - molecules. *Bio Systems*, *112*(2), 139–144. <https://doi.org/10.1016/j.biosystems.2013.03.004>
- Djordjevic, M., Djordjevic, M., & Severinov, K. (2012). CRISPR transcript processing: A mechanism for generating a large number of small interfering RNAs. *Biology Direct*, *7*, 24. <https://doi.org/10.1186/1745-6150-7-24>
- Dresch, J. M., Richards, M., & Ay, A. (2013). A primer on thermodynamic-based models for deciphering transcriptional regulatory logic. *Biochimica Et Biophysica Acta*, *1829*(9), 946–953. <https://doi.org/10.1016/j.bbagr.2013.04.011>
- Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, *403*(6767), 335–338. <https://doi.org/10.1038/35002125>
- Enikeeva, F. N., Severinov, K. V., & Gelfand, M. S. (2010). Restriction-modification systems and bacteriophage invasion: Who wins? *Journal of Theoretical Biology*, *266*(4), 550–559. <https://doi.org/10.1016/j.jtbi.2010.07.006>
- Ershova, A. S., Rusinov, I. S., Spirin, S. A., Karyagina, A. S., & Alexeevski, A. V. (2015). Role of Restriction-Modification Systems in Prokaryotic Evolution and Ecology. *Biochemistry. Biokhimiia*, *80*(10), 1373–1386. <https://doi.org/10.1134/S0006297915100193>
- Furuta, Y., & Kobayashi, I. (2012). Mobility of DNA sequence recognition domains in DNA methyltransferases suggests epigenetics-driven adaptive evolution. *Mobile Genetic Elements*, *2*(6), 292–296. <https://doi.org/10.4161/mge.23371>
- Furuta, Y., & Kobayashi, I. (2013). Restriction-Modification Systems as Mobile Epigenetic Elements. In *Madame Curie Bioscience Database [Internet]*. Landes Bioscience. <https://www.ncbi.nlm.nih.gov/books/NBK63963/>
- Gerosa, L., Kochanowski, K., Heinemann, M., & Sauer, U. (2013). Dissecting specific and global transcriptional regulation of bacterial gene expression. *Molecular Systems Biology*, *9*, 658. <https://doi.org/10.1038/msb.2013.14>
- Gertz, J., Siggia, E. D., & Cohen, B. A. (2009). Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature*, *457*(7226), 215–218. <https://doi.org/10.1038/nature07521>
- Gingeras, T. R., & Brooks, J. E. (1983). Cloned restriction/modification system from *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(2), 402–406. <https://doi.org/10.1073/pnas.80.2.402>
- Goldberg, G. W., & Marraffini, L. A. (2015). Resistance and tolerance to foreign elements by prokaryotic immune systems—Curating the genome. *Nature Reviews. Immunology*, *15*(11), 717–724. <https://doi.org/10.1038/nri3910>
- Graovac, S., Rodic, A., Djordjevic, M., Severinov, K., & Djordjevic, M. (2019). Effects of Population Dynamics on Establishment of a Restriction-Modification System in a Bacterial Host. *Molecules (Basel, Switzerland)*, *24*(1). <https://doi.org/10.3390/molecules24010198>
- Handa, N., Ichige, A., Kusano, K., & Kobayashi, I. (2000). Cellular responses to postsegregational killing by restriction-modification genes. *Journal of Bacteriology*, *182*(8), 2218–2229. <https://doi.org/10.1128/JB.182.8.2218-2229.2000>

- Handa, N., & Kobayashi, I. (1999). Post-segregational killing by restriction modification gene complexes: Observations of individual cell deaths. *Biochimie*, *81*(8–9), 931–938. [https://doi.org/10.1016/s0300-9084\(99\)00201-1](https://doi.org/10.1016/s0300-9084(99)00201-1)
- Hawley, D. K., & McClure, W. R. (1982). Mechanism of activation of transcription initiation from the lambda PRM promoter. *Journal of Molecular Biology*, *157*(3), 493–525. [https://doi.org/10.1016/0022-2836\(82\)90473-9](https://doi.org/10.1016/0022-2836(82)90473-9)
- He, X., Samee, M. A. H., Blatti, C., & Sinha, S. (2010). Thermodynamics-based models of transcriptional regulation by enhancers: The roles of synergistic activation, cooperative binding and short-range repression. *PLoS Computational Biology*, *6*(9), e1000935. <https://doi.org/10.1371/journal.pcbi.1000935>
- Hill, T. L. (2013). *Cooperativity Theory in Biochemistry: Steady-State and Equilibrium Systems*. Springer Science & Business Media.
- Hille, F., & Charpentier, E. (2016). CRISPR-Cas: Biology, mechanisms and relevance. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1707), 20150496. <https://doi.org/10.1098/rstb.2015.0496>
- Ichige, A., & Kobayashi, I. (2005). Stability of EcoRI restriction-modification enzymes in vivo differentiates the EcoRI restriction-modification system from other postsegregational cell killing systems. *Journal of Bacteriology*, *187*(19), 6612–6621. <https://doi.org/10.1128/JB.187.19.6612-6621.2005>
- Ingalls, B. P. (2013). *Mathematical Modeling in Systems Biology: An Introduction*. MIT Press.
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429–5433. <https://doi.org/10.1128/jb.169.12.5429-5433.1987>
- Ives, C. L., Nathan, P. D., & Brooks, J. E. (1992). Regulation of the BamHI restriction-modification system by a small intergenic open reading frame, bamHIC, in both Escherichia coli and Bacillus subtilis. *Journal of Bacteriology*, *174*(22), 7194–7201. <https://doi.org/10.1128/jb.174.22.7194-7201.1992>
- Jeltsch, A., & Pingoud, A. (1996). Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. *Journal of Molecular Evolution*, *42*(2), 91–96. <https://doi.org/10.1007/BF02198833>
- Karyagina, A., Shilov, I., Tashlitskii, V., Khodoun, M., Vasil'ev, S., Lau, P. C., & Nikolskaya, I. (1997). Specific binding of sso II DNA methyltransferase to its promoter region provides the regulation of sso II restriction-modification gene expression. *Nucleic Acids Research*, *25*(11), 2114–2120. <https://doi.org/10.1093/nar/25.11.2114>
- Katna, A., Boratynski, R., Furmanek-Blaszczak, B., Zolcinska, N., & Sektas, M. (2010). Unbalanced restriction impairs SOS-induced DNA repair effects. *Journal of Microbiology and Biotechnology*, *20*(1), 30–38.
- Kawamura, F., Mizukami, T., Shimotsu, H., Anzai, H., Takahashi, H., & Saito, H. (1981). Unusually infrequent cleavage with several endonucleases and physical map construction of Bacillus subtilis bacteriophage phi 1 DNA. *Journal of Virology*, *37*(3), 1099–1102. <https://doi.org/10.1128/JVI.37.3.1099-1102.1981>
- Kiss, A., Posfai, G., Keller, C. C., Venetianer, P., & Roberts, R. J. (1985). Nucleotide sequence of the BsuRI restriction-modification system. *Nucleic Acids Research*, *13*(18), 6403–6421. <https://doi.org/10.1093/nar/13.18.6403>

- Kita, K., Kotani, H., Sugisaki, H., & Takanami, M. (1989). The FokI restriction-modification system. I. Organization and nucleotide sequences of the restriction and modification genes. *The Journal of Biological Chemistry*, 264(10), 5751–5756.
- Kita, K., Tsuda, J., & Nakai, S. (2002). C.EcoO109I, a regulatory protein for production of EcoO109I restriction endonuclease, specifically binds to and bends DNA upstream of its translational start site. *Nucleic Acids Research*, 30(16), 3558–3565. <https://doi.org/10.1093/nar/gkf477>
- Kitano, H. (2007). A robustness-based approach to systems-oriented drug design. *Nature Reviews Drug Discovery*, 6(3), 202–210. <https://doi.org/10.1038/nrd2195>
- Klimuk, E., Bogdanova, E., Nagornykh, M., Rodic, A., Djordjevic, M., Medvedeva, S., Pavlova, O., & Severinov, K. (2018). Controller protein of restriction-modification system Kpn2I affects transcription of its gene by acting as a transcription elongation roadblock. *Nucleic Acids Research*, 46(20), 10810–10826. <https://doi.org/10.1093/nar/gky880>
- Klumpp, S., & Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition. *Current Opinion in Biotechnology*, 28, 96–102. <https://doi.org/10.1016/j.copbio.2014.01.001>
- Klumpp, S., Zhang, Z., & Hwa, T. (2009). Growth rate-dependent global effects on gene expression in bacteria. *Cell*, 139(7), 1366–1375. <https://doi.org/10.1016/j.cell.2009.12.001>
- Knowle, D., Lintner, R. E., Touma, Y. M., & Blumenthal, R. M. (2005). Nature of the promoter activated by C.PvuII, an unusual regulatory protein conserved among restriction-modification systems. *Journal of Bacteriology*, 187(2), 488–497. <https://doi.org/10.1128/JB.187.2.488-497.2005>
- Kobayashi, I. (2001). Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Research*, 29(18), 3742–3756. <https://doi.org/10.1093/nar/29.18.3742>
- Kobayashi, I. (2004). Restriction-Modification Systems as Minimal Forms of Life. In A. M. Pingoud (Ed.), *Restriction Endonucleases* (pp. 19–62). Springer. https://doi.org/10.1007/978-3-642-18851-0_2
- Korona, R., Korona, B., & Levin, B. R. (1993). Sensitivity of naturally occurring coliphages to type I and type II restriction and modification. *Journal of General Microbiology*, 139 Pt 6, 1283–1290. <https://doi.org/10.1099/00221287-139-6-1283>
- Korona, R., & Levin, B. R. (1993). Phage-mediated selection and the evolution and maintenance of restriction-modification. *Evolution; International Journal of Organic Evolution*, 47(2), 556–575. <https://doi.org/10.1111/j.1558-5646.1993.tb02113.x>
- Krüger, D. H., & Bickle, T. A. (1983). Bacteriophage survival: Multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiological Reviews*, 47(3), 345–360. <https://doi.org/10.1128/mr.47.3.345-360.1983>
- Krüger, D. H., & Schroeder, C. (1981). Bacteriophage T3 and bacteriophage T7 virus-host cell interactions. *Microbiological Reviews*, 45(1), 9–51. <https://doi.org/10.1128/mr.45.1.9-51.1981>
- Kulakauskas, S., Lubys, A., & Ehrlich, S. D. (1995). DNA restriction-modification systems mediate plasmid maintenance. *Journal of Bacteriology*, 177(12), 3451–3454. <https://doi.org/10.1128/jb.177.12.3451-3454.1995>
- Künne, T., Kieper, S. N., Bannenberg, J. W., Vogel, A. I. M., Miellet, W. R., Klein, M., Depken, M., Suarez-Diez, M., & Brouns, S. J. J. (2016). Cas3-Derived Target DNA Degradation Fragments Fuel Primed CRISPR Adaptation. *Molecular Cell*, 63(5), 852–864. <https://doi.org/10.1016/j.molcel.2016.07.011>

- Laursen, B. S., Sørensen, H. P., Mortensen, K. K., & Sperling-Petersen, H. U. (2005). Initiation of Protein Synthesis in Bacteria. *Microbiology and Molecular Biology Reviews*, 69(1), 101–123. <https://doi.org/10.1128/MMBR.69.1.101-123.2005>
- Le Novère, N. (2015). Quantitative and logic modelling of molecular and gene networks. *Nature Reviews. Genetics*, 16(3), 146–158. <https://doi.org/10.1038/nrg3885>
- Longo, D., & Hasty, J. (2006). Dynamics of single-cell gene expression. *Molecular Systems Biology*, 2, 64. <https://doi.org/10.1038/msb4100110>
- Lubys, A., & Janulaitis, A. (1995). Cloning and analysis of the plasmid-borne genes encoding the Bsp6I restriction and modification enzymes. *Gene*, 157(1–2), 25–29. [https://doi.org/10.1016/0378-1119\(94\)00795-t](https://doi.org/10.1016/0378-1119(94)00795-t)
- Lubys, A., Jurenaite, S., & Janulaitis, A. (1999). Structural organization and regulation of the plasmid-borne type II restriction-modification system Kpn2I from *Klebsiella pneumoniae* RFL2. *Nucleic Acids Research*, 27(21), 4228–4234.
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct*, 1, 7. <https://doi.org/10.1186/1745-6150-1-7>
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., ... Koonin, E. V. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews. Microbiology*, 13(11), 722–736. <https://doi.org/10.1038/nrmicro3569>
- Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2013). Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Research*, 41(8), 4360–4377. <https://doi.org/10.1093/nar/gkt157>
- Makarova, K. S., Wolf, Y. I., Snir, S., & Koonin, E. V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *Journal of Bacteriology*, 193(21), 6039–6056. <https://doi.org/10.1128/JB.05535-11>
- Makino, O., Saito, H., & Ando, T. (1980). *Bacillus subtilis*-phage phi 1 overcomes host-controlled restriction by producing BamNx inhibitor protein. *Molecular & General Genetics: MGG*, 179(3), 463–468. <https://doi.org/10.1007/BF00271735>
- McGeehan, J. E., Papapanagiotou, I., Streeter, S. D., & Kneale, G. G. (2006). Cooperative binding of the C.AhdI controller protein to the C/R promoter and its role in endonuclease gene expression. *Journal of Molecular Biology*, 358(2), 523–531. <https://doi.org/10.1016/j.jmb.2006.02.003>
- McGeehan, J. E., Streeter, S. D., Papapanagiotou, I., Fox, G. C., & Kneale, G. G. (2005). High-resolution crystal structure of the restriction-modification controller protein C.AhdI from *Aeromonas hydrophila*. *Journal of Molecular Biology*, 346(3), 689–701. <https://doi.org/10.1016/j.jmb.2004.12.025>
- Medina-Aparicio, L., Rebollar-Flores, J. E., Gallego-Hernández, A. L., Vázquez, A., Olvera, L., Gutiérrez-Ríos, R. M., Calva, E., & Hernández-Lucas, I. (2011). The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *Journal of Bacteriology*, 193(10), 2396–2407. <https://doi.org/10.1128/JB.01480-10>

- Mojica, F. J., Díez-Villaseñor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, *36*(1), 244–246. <https://doi.org/10.1046/j.1365-2958.2000.01838.x>
- Mojica, F. J. M., & Díez-Villaseñor, C. (2010). The on-off switch of CRISPR immunity against phages in *Escherichia coli*. *Molecular Microbiology*, *77*(6), 1341–1345. <https://doi.org/10.1111/j.1365-2958.2010.07326.x>
- Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, *60*(2), 174–182. <https://doi.org/10.1007/s00239-004-0046-3>
- Moll, I., Grill, S., Gualerzi, C. O., & Bläsi, U. (2002). Leaderless mRNAs in bacteria: Surprises in ribosomal recruitment and translational control. *Molecular Microbiology*, *43*(1), 239–246. <https://doi.org/10.1046/j.1365-2958.2002.02739.x>
- Morozova, N., Sabantsev, A., Bogdanova, E., Fedorova, Y., Maikova, A., Vedyaykin, A., Rodic, A., Djordjevic, M., Khodorkovskii, M., & Severinov, K. (2016). Temporal dynamics of methyltransferase and restriction endonuclease accumulation in individual cells after introducing a restriction-modification system. *Nucleic Acids Research*, *44*(2), 790–800. <https://doi.org/10.1093/nar/gkv1490>
- Morrison, M., Razo-Mejia, M., & Phillips, R. (2021). Reconciling kinetic and thermodynamic models of bacterial transcription. *PLoS Computational Biology*, *17*(1), e1008572. <https://doi.org/10.1371/journal.pcbi.1008572>
- Mruk, I., & Blumenthal, R. M. (2008). Real-time kinetics of restriction-modification gene expression after entry into a new host cell. *Nucleic Acids Research*, *36*(8), 2581–2593. <https://doi.org/10.1093/nar/gkn097>
- Mruk, I., & Kobayashi, I. (2014). To be or not to be: Regulation of restriction-modification systems and other toxin-antitoxin systems. *Nucleic Acids Research*, *42*(1), 70–86. <https://doi.org/10.1093/nar/gkt711>
- Mruk, I., Liu, Y., Ge, L., & Kobayashi, I. (2011). Antisense RNA associated with biological regulation of a restriction-modification system. *Nucleic Acids Research*, *39*(13), 5622–5632. <https://doi.org/10.1093/nar/gkr166>
- Mruk, I., Rajesh, P., & Blumenthal, R. M. (2007). Regulatory circuit based on autogenous activation-repression: Roles of C-boxes and spacer sequences in control of the PvuII restriction-modification system. *Nucleic Acids Research*, *35*(20), 6935–6952. <https://doi.org/10.1093/nar/gkm837>
- Musharova, O., Klimuk, E., Datsenko, K. A., Metlitskaya, A., Logacheva, M., Semenova, E., Severinov, K., & Savitskaya, E. (2017). Spacer-length DNA intermediates are associated with Cas1 in cells undergoing primed CRISPR adaptation. *Nucleic Acids Research*, *45*(6), 3297–3307. <https://doi.org/10.1093/nar/gkx097>
- Nagornykh, M. O., Bogdanova, E. S., Protsenko, A. S., Zakharova, M. V., Solonin, A. S., & Severinov, K. V. (2008). Regulation of gene expression in type II restriction-modification system. *Genetika*, *44*(5), 606–615. <https://doi.org/10.1134/s1022795408050037>
- Nagornykh, M., Zakharova, M., Protsenko, A., Bogdanova, E., Solonin, A. S., & Severinov, K. (2011). Regulation of gene expression in restriction-modification system Eco29kI. *Nucleic Acids Research*, *39*(11), 4653–4663. <https://doi.org/10.1093/nar/gkr055>
- Naito, T., Kusano, K., & Kobayashi, I. (1995). Selfish behavior of restriction-modification systems. *Science (New York, N.Y.)*, *267*(5199), 897–899. <https://doi.org/10.1126/science.7846533>
- Neuhauser, C. (2011). *Calculus For Biology and Medicine* (3rd edition). Pearson Education.

- O'Donnell, S. M., & Janssen, G. R. (2001). The initiation codon affects ribosome binding and translational efficiency in *Escherichia coli* of cI mRNA with or without the 5' untranslated leader. *Journal of Bacteriology*, *183*(4), 1277–1283. <https://doi.org/10.1128/JB.183.4.1277-1283.2001>
- Ohno, S., Handa, N., Watanabe-Matsui, M., Takahashi, N., & Kobayashi, I. (2008). Maintenance forced by a restriction-modification system can be modulated by a region in its modification enzyme not essential for methyltransferase activity. *Journal of Bacteriology*, *190*(6), 2039–2049. <https://doi.org/10.1128/JB.01319-07>
- Patterson, A. G., Yevstigneyeva, M. S., & Fineran, P. C. (2017). Regulation of CRISPR-Cas adaptive immune systems. *Current Opinion in Microbiology*, *37*, 1–7. <https://doi.org/10.1016/j.mib.2017.02.004>
- Phillips, R., Kondev, J., Theriot, J., Garcia, H. G., & Orme, N. (2012). *Physical Biology of the Cell* (2nd ed.). Garland Science. <https://doi.org/10.1201/9781134111589>
- Pickar-Oliver, A., & Gersbach, C. A. (2019). The next generation of CRISPR-Cas technologies and applications. *Nature Reviews. Molecular Cell Biology*, *20*(8), 490–507. <https://doi.org/10.1038/s41580-019-0131-5>
- Pingoud, A., Fuxreiter, M., Pingoud, V., & Wende, W. (2005). Type II restriction endonucleases: Structure and mechanism. *Cellular and Molecular Life Sciences: CMLS*, *62*(6), 685–707. <https://doi.org/10.1007/s00018-004-4513-1>
- Pingoud, A., Wilson, G. G., & Wende, W. (2014). Type II restriction endonucleases—A historical perspective and more. *Nucleic Acids Research*, *42*(12), 7489–7527. <https://doi.org/10.1093/nar/gku447>
- Porreca, R., Cinquemani, E., Lygeros, J., & Ferrari-Trecate, G. (2010). Structural identification of unate-like genetic network models from time-lapse protein concentration measurements. *49th IEEE Conference on Decision and Control (CDC)*, 2529–2534. <https://doi.org/10.1109/CDC.2010.5717922>
- Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K. A., Djordjevic, M., Wanner, B. L., & Severinov, K. (2010). Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Molecular Microbiology*, *77*(6), 1367–1379. <https://doi.org/10.1111/j.1365-2958.2010.07265.x>
- Prakash-Cheng, A., Chung, S. S., & Ryu, J. (1993). The expression and regulation of hsdK genes after conjugative transfer. *Molecular & General Genetics: MGG*, *241*(5–6), 491–496. <https://doi.org/10.1007/BF00279890>
- Prakash-Cheng, A., & Ryu, J. (1993). Delayed expression of in vivo restriction activity following conjugal transfer of *Escherichia coli* hsdK (restriction-modification) genes. *Journal of Bacteriology*, *175*(15), 4905–4906. <https://doi.org/10.1128/jb.175.15.4905-4906.1993>
- Price, C., & Bickle, T. A. (1986). A possible role for DNA restriction in bacterial evolution. *Microbiological Sciences*, *3*(10), 296–299.
- Price, V. J., Huo, W., Sharifi, A., & Palmer, K. L. (2016). CRISPR-Cas and Restriction-Modification Act Additively against Conjugative Antibiotic Resistance Plasmid Transfer in *Enterococcus faecalis*. *MSphere*, *1*(3). <https://doi.org/10.1128/mSphere.00064-16>
- Protsenko, A., Zakharova, M., Nagornykh, M., Solonin, A., & Severinov, K. (2009). Transcription regulation of restriction-modification system Ecl18kI. *Nucleic Acids Research*, *37*(16), 5322–5330. <https://doi.org/10.1093/nar/gkp579>
- Ptashne, M. (2005). Regulation of transcription: From lambda to eukaryotes. *Trends in Biochemical Sciences*, *30*(6), 275–279. <https://doi.org/10.1016/j.tibs.2005.04.003>

- Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., & Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Molecular Microbiology*, *75*(6), 1495–1512. <https://doi.org/10.1111/j.1365-2958.2010.07073.x>
- Ratner, H. K., Sampson, T. R., & Weiss, D. S. (2015). I can see CRISPR now, even when phage are gone: A view on alternative CRISPR-Cas functions from the prokaryotic envelope. *Current Opinion in Infectious Diseases*, *28*(3), 267–274. <https://doi.org/10.1097/QCO.000000000000154>
- Rezulak, M., Borsuk, I., & Mruk, I. (2016). Natural C-independent expression of restriction endonuclease in a C protein-associated restriction-modification system. *Nucleic Acids Research*, *44*(6), 2646–2660. <https://doi.org/10.1093/nar/gkv1331>
- Richter, C., Chang, J. T., & Fineran, P. C. (2012). Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. *Viruses*, *4*(10), 2291–2311. <https://doi.org/10.3390/v4102291>
- Rimseliene, R., Vaisvila, R., & Janulaitis, A. (1995). The *eco72IC* gene specifies a trans-acting factor which influences expression of both DNA methyltransferase and endonuclease from the *Eco72I* restriction-modification system. *Gene*, *157*(1–2), 217–219. [https://doi.org/10.1016/0378-1119\(94\)00794-s](https://doi.org/10.1016/0378-1119(94)00794-s)
- Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. (2005). REBASE--restriction enzymes and DNA methyltransferases. *Nucleic Acids Research*, *33*(Database issue), D230-232. <https://doi.org/10.1093/nar/gki029>
- Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. (2015). REBASE--a database for DNA restriction and modification: Enzymes, genes and genomes. *Nucleic Acids Research*, *43*(Database issue), D298-299. <https://doi.org/10.1093/nar/gku1046>
- Rocha, E. P., Danchin, A., & Viari, A. (2001). Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. *Genome Research*, *11*(6), 946–958. <https://doi.org/10.1101/gr.gr-1531rr>
- Rodic, A., Blagojevic, B., & Djordjevic, M. (2018). Systems Biology of Bacterial Immune Systems: Regulation of Restriction-Modification and CRISPR-Cas Systems. In N. Rajewsky, S. Jurga, & J. Barciszewski (Eds.), *Systems Biology* (pp. 37–58). Springer International Publishing. https://doi.org/10.1007/978-3-319-92967-5_3
- Rodic, A., Blagojevic, B., Djordjevic, M., Severinov, K., & Djordjevic, M. (2017). Features of CRISPR-Cas Regulation Key to Highly Efficient and Temporally-Specific crRNA Production. *Frontiers in Microbiology*, *8*, 2139. <https://doi.org/10.3389/fmicb.2017.02139>
- Rodic, A., Blagojevic, B., Zdobnov, E., Djordjevic, M., & Djordjevic, M. (2017). Understanding key features of bacterial restriction-modification systems through quantitative modeling. *BMC Systems Biology*, *11*(Suppl 1), 377. <https://doi.org/10.1186/s12918-016-0377-x>
- Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., & Elowitz, M. B. (2005). Gene regulation at the single-cell level. *Science (New York, N.Y.)*, *307*(5717), 1962–1965. <https://doi.org/10.1126/science.1106914>
- Savageau, M. A. (1977). Design of molecular control mechanisms and the demand for gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5647–5651. <https://doi.org/10.1073/pnas.74.12.5647>
- Sawaya, M. R., Zhu, Z., Mersha, F., Chan, S.-H., Dabur, R., Xu, S.-Y., & Balendiran, G. K. (2005). Crystal structure of the restriction-modification system control element C.BclI and mapping of its binding site. *Structure (London, England: 1993)*, *13*(12), 1837–1847. <https://doi.org/10.1016/j.str.2005.08.017>

- Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U., & Gaul, U. (2008). Predicting expression patterns from regulatory sequence in *Drosophila* segmentation. *Nature*, *451*(7178), 535–540. <https://doi.org/10.1038/nature06496>
- Semenova, E., Minakhin, L., Bogdanova, E., Nagornykh, M., Vasilov, A., Heyduk, T., Solonin, A., Zakharova, M., & Severinov, K. (2005). Transcription regulation of the EcoRV restriction-modification system. *Nucleic Acids Research*, *33*(21), 6942–6951. <https://doi.org/10.1093/nar/gki998>
- Severinov, K., Ispolatov, I., & Semenova, E. (2016). The Influence of Copy-Number of Targeted Extrachromosomal Genetic Elements on the Outcome of CRISPR-Cas Defense. *Frontiers in Molecular Biosciences*, *3*, 45. <https://doi.org/10.3389/fmolb.2016.00045>
- Shea, M. A., & Ackers, G. K. (1985). The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. *Journal of Molecular Biology*, *181*(2), 211–230. [https://doi.org/10.1016/0022-2836\(85\)90086-5](https://doi.org/10.1016/0022-2836(85)90086-5)
- Sneppen, K., & Zocchi, G. (2005). *Physics in Molecular Biology*. Cambridge University Press.
- Sohail, A., Ives, C. L., & Brooks, J. E. (1995). Purification and characterization of C.BamHI, a regulator of the BamHI restriction-modification system. *Gene*, *157*(1–2), 227–228. [https://doi.org/10.1016/0378-1119\(94\)00698-r](https://doi.org/10.1016/0378-1119(94)00698-r)
- Som, S., & Friedman, S. (1994). Regulation of EcoRII methyltransferase: Effect of mutations on gene expression and in vitro binding to the promoter region. *Nucleic Acids Research*, *22*(24), 5347–5353. <https://doi.org/10.1093/nar/22.24.5347>
- Som, S., & Friedman, S. (1997). Characterization of the intergenic region which regulates the MspI restriction-modification system. *Journal of Bacteriology*, *179*(3), 964–967. <https://doi.org/10.1128/jb.179.3.964-967.1997>
- Sorokin, V., Severinov, K., & Gelfand, M. S. (2009). Systematic prediction of control proteins and their DNA binding sites. *Nucleic Acids Research*, *37*(2), 441–451. <https://doi.org/10.1093/nar/gkn931>
- Sorokin, V., Severinov, K., & Gelfand, M. S. (2010). Large-scale identification and analysis of C-proteins. *Methods in Molecular Biology (Clifton, N.J.)*, *674*, 269–282. https://doi.org/10.1007/978-1-60761-854-6_17
- Srihanta, Y. N., Gorrell, R. J., Steen, J. A., Gawthorne, J. A., Kwok, T., Grimmond, S. M., Robins-Browne, R. M., & Jennings, M. P. (2011). Phasevarion mediated epigenetic gene regulation in *Helicobacter pylori*. *PLoS One*, *6*(12), e27569. <https://doi.org/10.1371/journal.pone.0027569>
- Stayrook, S., Jaru-Ampornpan, P., Ni, J., Hochschild, A., & Lewis, M. (2008). Crystal structure of the lambda repressor and a model for pairwise cooperative operator binding. *Nature*, *452*(7190), 1022–1025. <https://doi.org/10.1038/nature06831>
- Stefan, D., Pinel, C., Pinhal, S., Cinquemani, E., Geiselmann, J., & de Jong, H. (2015). Inference of quantitative models of bacterial promoters from time-series reporter gene data. *PLoS Computational Biology*, *11*(1), e1004028. <https://doi.org/10.1371/journal.pcbi.1004028>
- Sternberg, S. H., Richter, H., Charpentier, E., & Qimron, U. (2016). Adaptation in CRISPR-Cas Systems. *Molecular Cell*, *61*(6), 797–808. <https://doi.org/10.1016/j.molcel.2016.01.030>
- Stowe, K. (2007). *An Introduction to Thermodynamics and Statistical Mechanics*. Cambridge University Press.
- Stratmann, T., Pul, Ü., Wurm, R., Wagner, R., & Schnetz, K. (2012). RcsB-BglJ activates the *Escherichia coli* leuO gene, encoding an H-NS antagonist and pleiotropic regulator of

- virulence determinants. *Molecular Microbiology*, 83(6), 1109–1123. <https://doi.org/10.1111/j.1365-2958.2012.07993.x>
- Streeter, S. D., Papapanagiotou, I., McGeehan, J. E., & Kneale, G. G. (2004). DNA footprinting and biophysical characterization of the controller protein C.AhdI suggests the basis of a genetic switch. *Nucleic Acids Research*, 32(21), 6445–6453. <https://doi.org/10.1093/nar/gkh975>
- Takahashi, I., & Marmur, J. (1963). Replacement of thymidylic acid by deoxyuridylic acid in the deoxyribonucleic acid of a transducing phage for *Bacillus subtilis*. *Nature*, 197, 794–795. <https://doi.org/10.1038/197794a0>
- Tao, T., Bourne, J. C., & Blumenthal, R. M. (1991). A family of regulatory genes associated with type II restriction-modification systems. *Journal of Bacteriology*, 173(4), 1367–1375. <https://doi.org/10.1128/jb.173.4.1367-1375.1991>
- Tock, M. R., & Dryden, D. T. F. (2005). The biology of restriction and anti-restriction. *Current Opinion in Microbiology*, 8(4), 466–472. <https://doi.org/10.1016/j.mib.2005.06.003>
- Touchon, M., Charpentier, S., Clermont, O., Rocha, E. P. C., Denamur, E., & Branger, C. (2011). CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *Journal of Bacteriology*, 193(10), 2460–2467. <https://doi.org/10.1128/JB.01307-10>
- Vasu, K., & Nagaraja, V. (2013). Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiology and Molecular Biology Reviews: MMBR*, 77(1), 53–72. <https://doi.org/10.1128/MMBR.00044-12>
- Venturelli, O. S., Egbert, R. G., & Arkin, A. P. (2016). Towards Engineering Biological Systems in a Broader Context. *Journal of Molecular Biology*, 428(5 Pt B), 928–944. <https://doi.org/10.1016/j.jmb.2015.10.025>
- Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A. C., Engelberg-Kulka, H., & Moll, I. (2011). Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell*, 147(1), 147–157. <https://doi.org/10.1016/j.cell.2011.07.047>
- Vijesurier, R. M., Carlock, L., Blumenthal, R. M., & Dunbar, J. C. (2000). Role and mechanism of action of C. PvuII, a regulatory protein conserved among restriction-modification systems. *Journal of Bacteriology*, 182(2), 477–487. <https://doi.org/10.1128/JB.182.2.477-487.2000>
- Vilar, J. M. G., Guet, C. C., & Leibler, S. (2003). Modeling network dynamics: The lac operon, a case study. *The Journal of Cell Biology*, 161(3), 471–476. <https://doi.org/10.1083/jcb.200301125>
- Vilar, J. M. G., & Leibler, S. (2003). DNA looping and physical constraints on transcription regulation. *Journal of Molecular Biology*, 331(5), 981–989. [https://doi.org/10.1016/s0022-2836\(03\)00764-2](https://doi.org/10.1016/s0022-2836(03)00764-2)
- Wang, L.-Z., Wu, F., Flores, K., Lai, Y.-C., & Wang, X. (2016). Build to understand: Synthetic approaches to biology. *Integrative Biology: Quantitative Biosciences from Nano to Macro*, 8(4), 394–408. <https://doi.org/10.1039/c5ib00252d>
- Westra, E. R., Pul, U., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M., Van Heereveld, L., Mastop, M., Wagner, E. G. H., Schnetz, K., Van Der Oost, J., Wagner, R., & Brouns, S. J. J. (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Molecular Microbiology*, 77(6), 1380–1393. <https://doi.org/10.1111/j.1365-2958.2010.07315.x>

- Williams, K., Savageau, M. A., & Blumenthal, R. M. (2013). A bistable hysteretic switch in an activator-repressor regulated restriction-modification system. *Nucleic Acids Research*, *41*(12), 6045–6057. <https://doi.org/10.1093/nar/gkt324>
- Wilson, G. G. (1991). Organization of restriction-modification systems. *Nucleic Acids Research*, *19*(10), 2539–2566. <https://doi.org/10.1093/nar/19.10.2539>
- Wilson, G. G., & Murray, N. E. (1991). Restriction and modification systems. *Annual Review of Genetics*, *25*, 585–627. <https://doi.org/10.1146/annurev.ge.25.120191.003101>
- Yang, C.-D., Chen, Y.-H., Huang, H.-Y., Huang, H.-D., & Tseng, C.-P. (2014). CRP represses the CRISPR/Cas system in *Escherichia coli*: Evidence that endogenous CRISPR spacers impede phage P1 replication. *Molecular Microbiology*, *92*(5), 1072–1091. <https://doi.org/10.1111/mmi.12614>
- Young, J. W., Locke, J. C. W., Altinok, A., Rosenfeld, N., Bacarian, T., Swain, P. S., Mjolsness, E., & Elowitz, M. B. (2011). Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy. *Nature Protocols*, *7*(1), 80–88. <https://doi.org/10.1038/nprot.2011.432>
- Zakharova, M., Minakhin, L., Solonin, A., & Severinov, K. (2004). Regulation of RNA polymerase promoter selectivity by covalent modification of DNA. *Journal of Molecular Biology*, *335*(1), 103–111. <https://doi.org/10.1016/j.jmb.2003.09.081>
- Zhang, C., Tsoi, R., & You, L. (2016). Addressing biological uncertainties in engineering gene circuits. *Integrative Biology: Quantitative Biosciences from Nano to Macro*, *8*(4), 456–464. <https://doi.org/10.1039/c5ib00275c>
- Zheleznaya, L. A., Kainov, D. E., Yunusova, A. K., & Matvienko, N. I. (2003). Regulatory C protein of the EcoRV modification-restriction system. *Biochemistry. Biokhimiia*, *68*(1), 105–110. <https://doi.org/10.1023/a:1022105804578>

Биографија аутора

Анђела М. Родић, рођена 19.9.1991. у Београду, је завршила основне академске студије Биологије (модул Молекуларна биологија и физиологија) и мастер академске студије Молекуларне биологије и физиологије (модул Биофизика) на Биолошком факултету Универзитета у Београду, да би 2015. уписала мултидисциплинарне докторске академске студије Биофизике на Универзитету у Београду. Од 2015. је запослена на Биолошком факултету, при чему од 2017. ради у наставном звању асистента на групи предмета из биофизике.

Осим држањем наставе, до сада се примарно бавила научним истраживањем из области молекуларне системске биологије, на 4 државна и једном међународном пројекту. Објекат овог истраживања су били молекуларни системи који бране бактерије од унете ДНК чије је регулаторне механизме и принципе дизајна проучавала путем конструисања и анализирања математичких модела. Добијени резултати су објављени у оквиру 5 оригиналних научних радова средњег IF 6.3 (енг. *Impact Factor*) и једног поглавља у књизи издавача *Springer Nature*. У додатних 5 радова средњег IF 4.2 су објављени резултати њеног новијег истраживања у коме је путем математичког моделовања и статистичких метода испитана зависност развоја епидемије COVID-19 од фактора средине и популације.

У периоду основних студија, Анђела је примљена на студентске праксе на Институту за нуклеарне науке „Винча“ и Универзитету у Лођу, Пољска. Фондација „Горан Љубијанкић“ јој је доделила годишњу награду за најбољи мастер рад из области молекуларне биологије. Анђела је одржала презентације на више међународних и домаћих научних скупова, међу којима и два предавања по позиву. Добитник је стипендија за учешће у неколико међународних школа биофизике (у Аустрији, Хрватској и Бразилу), а на једној је освојила и награду за најбољу презентацију (у Русији).

Изјава о ауторству

Име и презиме аутора: Анђела Родић

Број индекса: 123/2015

Изјављујем

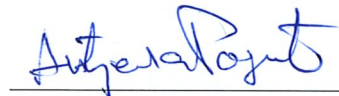
да је докторска дисертација под насловом

Биофизичко моделовање рестрикционо-модификационих система бактерија

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

У Београду, 28.03.2022.

Потпис аутора



Anđela Rodić

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора: Анђела Родић

Број индекса: 123/2015

Студијски програм: Биофизика

Наслов рада: Биофизичко моделовање рестрикционо-модификационих система бактерија

Ментори:

- Проф. др Марко Ђорђевић, ванредни професор, Универзитет у Београду – Биолошки факултет
- Др Магдалена Ђорђевић, научни саветник, Универзитет у Београду – Институт за физику

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у **Дигиталном репозиторијуму Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

У Београду, 28.03.2022.

Потпис аутора



Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Биофизичко моделовање рестрикционо-модификационих система бактерија

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.

Кратак опис лиценци је саставни део ове изјаве).

У Београду, 28.03.2022.

Потпис аутора