



Transzkripció időzítés és genetikai hálózatelemzés az
Escherichia coli galaktóz hasznosítási hálózatán

– doktori értekezés –

Horváth Péter

Témavezető: Dr. Semsey Szabolcs egyetemi adjunktus

ELTE TTK Biológia Doktori iskola

Iskolavezető: Dr. Erdei Anna egyetemi tanár, az MTA r. tagja.

Klasszikus és Molekuláris Genetika program

Programvezető: Dr. Orosz László egyetemi tanár, az MTA r. tagja.

Budapest, 2010.

Tartalomjegyzék

Tartalomjegyzék	2
Rövidítések jegyzéke	4
1. Bevezetés	5
2. Irodalmi áttekintés	8
2.1 Szabályozási motívumok a kis molekulákat hasznosító rendszerekben	8
2.2 Az <i>E. coli</i> galaktóz hasznosításért felelős hálózata	10
2.2.1 A szabályozó gének	11
2.2.2 A cAMP-CRP komplex	15
2.2.3 Szabályozó régiók	17
2.2.4 Transzport gének és szabályozásuk	18
2.2.5 Metabolizmus gének és szabályozásuk	21
2.3 A Boole-algebra, mint a szabályozás leírásának kvalitatív módja	28
2.4 A transzkripció szabályozás matematikai leírásának alapjai	30
3. Célkitűzés	34
4. Anyagok és módszerek	35
4.1 Törzsek	35
4.2 Táploldatok és kiegészítőik	36
4.3 Klónozás	37
4.4 Promóter konstrukciók létrehozása	38
4.5 Riportertörzsek létrehozása	39
4.6 Elektrokompetens sejt készítése	40
4.7 A szignál integráció vizsgálata <i>in vivo</i>	40
4.8 Törzsek ellenőrzése szekvenálással	41
4.9 β -glükuronidáz assay	41
4.10 <i>In vitro</i> transzkripció	42
4.11 A <i>gal</i> rendszer promóter aktivitásainak modellje	43
4.12 A D-galaktóz és a cAMP-CRP szintek meghatározása	46
5. Eredmények	48
5.1 A <i>gal</i> rendszer időzítéséhez kapcsolódó eredmények	48

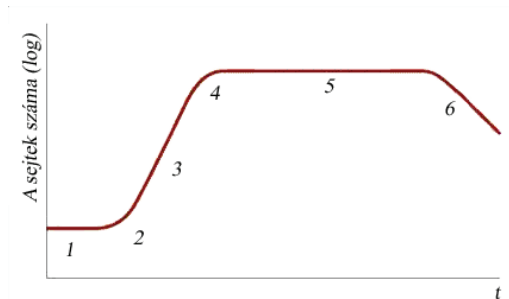
5.1.1 A növekedési görbe kiszámítása	49
5.1.2 A promóter aktivitások kiszámítása	50
5.1.3 Az intracelluláris cAMP és D-galaktóz szint számítása	52
5.1.4 Változások a $P1_{galE}$ és a $P2_{galE}$ transzkripciójában	53
5.1.5 A gal regulon promótereinek transzkripciója a stacionárius fázisra jellemző (σ^{38}) RNS-polimerázzal	54
5.2 A gal rendszer genetikai flexibilitásával kapcsolatos eredmények	55
5.2.1 Logikai kapuk kombinálása	59
6. Diskusszió	60
6.1 A galaktóz transzport és a felhasználás koordinálása	60
6.2 A galaktóz rendszer promótereinek transzkripció mintázatát meghatározó tényezők	62
6.3 A szabályozási hálózat flexibilitása	63
6.4 A logikai kapuk kombinálása	64
7. Összefoglalás	66
8 Summary	68
9. Köszönetnyilvánítás	70
10. Felhasznált irodalom	71

Rövidítések jegyzéke

- α CTD – Az RNS polimeráz α alegységének C-terminális doménje.
- cAMP-CRP – Ciklikus adenzin-monofoszfát és receptor fehérjéjének (CRP) komplexe
- CAP – Catabolite Activator Protein (katabolit aktivátor fehérje), a CRP alternatív neve
- CDTA – Ciklohexán-1,2-diamino-tetraecetsav
- CRP – cAMP Receptor Protein
- DTT – Ditiotritol, redukáló szer
- EIIA^{Gk} – A PTS rendszer Enzyme II (EII) fehérjéjének „A” doménje.
- GalR – Galaktóz represszor
- GalS – Galaktóz izorepresszor
- GBP – Galaktózkötő fehérje (Galactose binding Protein)
- GusA – β -glükuronidáz
- HU – Heat Unstable Nucleoid Protein
- LacI – laktóz represszor
- LB – Luria-Bertani táptalaj
- O_E – A *galETKM* operon külső GalR-operátora.
- O_I – A *galETKM* operon belső GalR-operátora.
- PNPG – para-nitrofenil- β -D-galaktopiranozid
- RNAP – RNS-polimeráz
- PTS – foszfoenol-piruvát-szénhidrát foszfotransferáz rendszer
- TF – Transzkripció faktor
- X-Gluc – 5-bromo-4-kloro-3-indolil- β -D-glükuronsav

1. Bevezetés

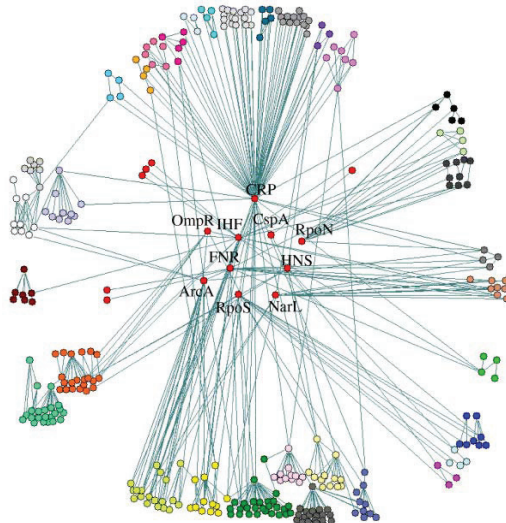
A szabadon élő baktériumsejtek növekedését – aminek modellje pl. a folyadékban növesztett *Escherichia coli* – leíró paradigma ma az ún. feast-famine ciklus, ami azt jelenti, hogy a sejtek számára nem áll rendelkezésre állandóan a táplálék, hanem időről-időre egy limitált mennyiségű tápanyaghoz jutnak, azon növekedésnek indulnak, majd a forrás kifogytával a növekedésük lelassul, megáll[1]. A folyadékkultúrában növesztett baktériumoknak, amik egy ilyen feast-famine ciklus modelljeként szolgálhatnak, hat különböző növekedési állapotát különböztethetjük meg (1.1 ábra): lag fázis (a növekedési ráta nulla), gyorsuló fázis (a növekedési ráta növekszik), exponenciális fázis (a növekedési ráta állandó), lassuló fázis (a növekedési ráta csökken) stacioner fázis (a növekedési ráta nulla) és a hanyatlás fázisa (a növekedési ráta negatív)[2]. Az ezeket a ciklusokat kísérő morfológiai és metabolikus változásokat korábban részletesen tanulmányozták[1].



1.1 ábra. A folyadékkultúrában növekvő baktériumok növekedési fázisai. 1.: Lag fázis, 2.: gyorsuló fázis, 3.: exponenciális fázis, 4.: lassuló fázis, 5.: stacioner fázis, 6.: hanyatlás.

A sejtnek érzékelnie kell a külső és belső környezet változásait, ezekre valamilyen választ kell adnia, illetve szabályoznia kell azt, hogy ezen válaszok számos alfolyamata koordináltan játszódjon le, azaz a különböző géntermékek a megfelelő időben és megfelelő arányban keletkezzenek.

Egy *E. coli* sejtben egyidejűleg körülbelül 1850-féle fehérjemolekula található[3], az *E. coli* genomon mai napig azonosított gének száma ugyanakkor 4300 fölött van[4]. Némelyik fehérjéből igen nagy mennyiség szintetizálódik, másokból alig pár darab. A fehérjék többnyire nem egymagukban teszik a dolgukat, hanem valamilyen kölcsönható partnerrel, amivel meg kell találniuk egymást. E kölcsönhatások közül ma már nagyon sokat ismerünk, és ha a kölcsönhatásokat egy hálózatként próbáljuk értelmezni, akkor egy igen összetett hálózatot kapunk. (1.2 ábra)



1.2 ábra. Az *E. coli* transzkripció szabályozási hálózatának funkcionális moduljai és azok kapcsolatai. A színes körök operonokat jelölnek, az azonos színűek egyazon funkcionális modulba tartoznak. A tíz legfontosabbnak tartott központi szabályozó faktor nevét feltünteti az ábra[5]. Ezek közül a CRP (cAMP Receptor Protein) jelen dolgozat témájában is fontos szerepet játszik.

Mindebből nagyon jól látszik, hogy a sejtnek valamilyen módon el kell döntenie, hogy mikor melyik génről írjon át génterméket, nem teheti meg, hogy minden génjét mindig aktívan tartja, vagyis a válaszait valamilyen szempontból optimalizálnia kell. Optimalizációs szempont lehet például a gyors reagálás képessége akár költségek árán, vagy a minimális energia befektetés szükségessége.

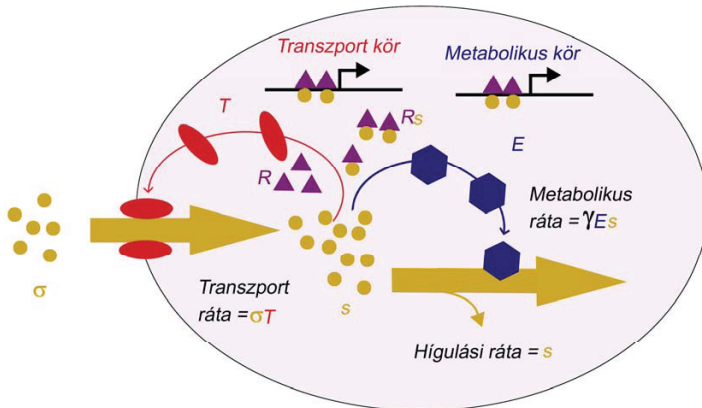
Az *E. coli* különböző alrendszerait irányító folyamatok elemei közti kölcsönhatások leírásával, illetve azok mennyiségi jellemzésével egyre komplexebb hálózatok rajzolódnak ki a szemünk előtt. Ugyanakkor a hálózat feltérképezése (vagyis a kölcsönható partnerek felírása és a kölcsönhatás típusának megállapítása) még nem ad támpontot arra nézve, hogy egy dinamikus környezetben hogyan viselkedik a hálózat, illetve egyes tipikus hálózati minták, motívumok milyen szempontból optimalizáltak. Ezek megértésében a laborkísérleteken túl a matematikai modellezés eszköztára is szükség van.

Jelen dolgozat az *E. coli* galaktóz hasznosítási hálózatán végzett kutatásainkat mutatja be. A munka célja egy részről az volt, hogy feltárja a hálózat tagjainak időzítési mintázatait egy modellezett feast-famine ciklus során, más részről hogy a hálózat szabályozóelemeinek logikai működésére, a bennük rejlő evolúciós potenciálra próbáljon fényt deríteni.

2. Irodalmi áttekintés

2.1 Szabályozási motívumok a kis molekulákat hasznosító rendszerekben

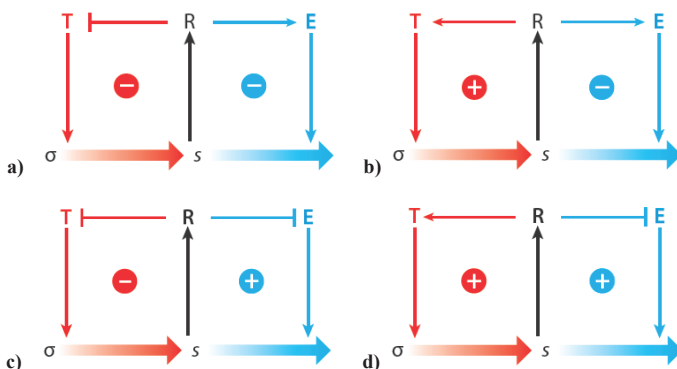
A sejteket irányító mechanizmusok megismerésére kétféle stratégia adódik: felülről és alulról építkező, melyek közül mi az alulról építkező stratégiát követjük. Ez a megközelítés alhálózatok kvantitatív leírásával és alapvető motívumok, visszacsatolási körök megértésével törekszik a teljes hálózat megismerésére. Ilyen jól leírt hálózatok többek között a kismolekulák, például cukrok felhasználásáért felelős hálózatok, melyek általános felépítése a 2.1 ábrán látható.



2.1 ábra. Egy kis molekula felhasználási hálózatának általános sémája. A lila ovális terület a baktérium. A kismolekula extracelluláris koncentrációja σ , intracelluláris koncentrációja s . A kis molekula a sejtbe az extracelluláris koncentráció és a transzport fehérjék koncentrációja (T) által meghatározott rátával jut be, és az aktuális intracelluláris koncentrációnak és a metabolikus enzimek koncentrációjának (E) függvényében használódik fel, valamint a sejtosztódás miatti hígulással csökkenhet a koncentrációja. A kismolekula sejtben belüli koncentrációját R szabályozó fehérje érzékeli, ez szabályozza a transzportot és a metabolizmust. A metabolikus ráta esetén γ azt a rátát jelöli, amivel a sejt a kismolekulát metabolizálja ahhoz képest, ahogy a kismolekula az osztódás miatt hígul[6].

Az ilyen rendszer két visszacsatolási köre épül, melyeket egy közös szabályozó elem kapcsol össze. Ez érzékeli a kis molekula intracelluláris szintjét [7]. A legegyszerűbb rendszerek, mint az *E. coli* laktóz hasznosítási rendszere két operonból áll, egy regulátor génből, és az általa szabályozott operonból, ami legalább két cisztront tartalmaz: egy transzporter fehérjét kódoló, és egy, a kismolekulát lebontó vagy módosító enzimet kódoló cisztront[8].

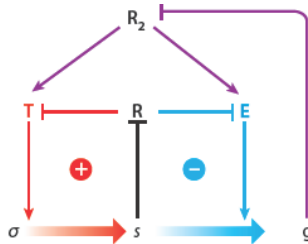
Két visszacsatolási kört elméletileg négyféle módon lehet összekapcsolni annak függvényében, hogy melyik pozitív és melyik negatív (2.2 ábra), bár ezek közül baktériumokban eddig csak kettőre sikerült példát találni[7].



2.2 ábra. A kis molekulák hasznosításáért felelős kettős visszacsatolási kör négy lehetséges megoldása. T, R és E rendre a transzporter, a regulátor és a hasznosításért felelős enzim mennyiségét jelöli, σ és s a kismolekula extracelluláris és intracelluláris szintjét jelenti, a vastag nyíl pedig ennek áramlása a rendszeren keresztül. (A vékony nyilak és talpas nyilak az aktiváció és a gátlás szokásos értelmében szerepelnek.) A negatív és a pozitív jelzés magára a kismolekula intracelluláris szintjére utal, azaz a transzportkör (piros) esetén a + fokozott transzportot, ezáltal s felhalmozódást jelent, a felhasználás (kék) esetén a + csökkent felhasználást, ezáltal szintén s felhalmozódást jelent. Az a) motívum működése során az intracelluláris kismolekula koncentráció szűk határokon belül mozog (homeosztázis): alacsony s esetén a transzport kör kapcsol be és a felhasználás szünetel, magas s esetén fordítva. Ilyen motívum szabályozza a sejtek vas anyagcseréjét. A b) motívum a sejten kívüli kismolekula transzportját és felhasználását teszi lehetővé (azaz ha megjelenik a kismolekula, akkor mindkét rendszer bekapcsol). Tipikus példa az *E. coli* cukorhasznosítási rendszerei közül a *lac* vagy a *gal* rendszer. A c) motívum az előző fordítottja. Működése a divatcikkek világára hasonlít, azaz minél több van egy divatcikkből a piacon, annál kevésbé kelendő (csökken a felhasználás). A d) motívum esetében s egyrészt gátolja önmaga felhasználását, másrészt további transzportot indukál, így önmaga felhalmozódását serkenti[7].

2.2 Az *E. coli* galaktóz hasznosításért felelős hálózata

Az *E. coli* cukorhasznosító rendszereinek alapsémája a fogyasztó motívum (2.2 b) ábra). Ez a séma kiegészíthető további elemekkel, például egy további szabályozó elemmel, ami valamiképpen módosítja a motívum eredeti működését (2.3 ábra).



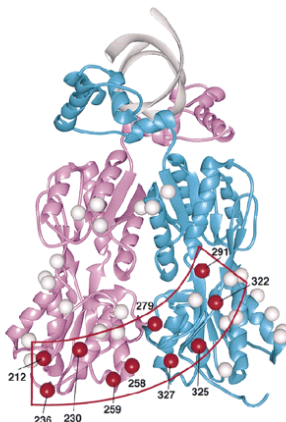
2.3 ábra. „Szükség esetén fogyasztó” motívum. T, R és E rendre a transzpozterer, a regulátor és a hasznosításért felelős enzim mennyiségét jelöli, σ és s a kismolekula extracelluláris és intracelluláris szintjét jelenti, σ és s a kismolekula extracelluláris és intracelluláris szintjét jelenti, a vastag nyíl pedig ennek áramlásba a rendszeren keresztül. R_2 egy második regulátor. (A vékony nyílak és talpas nyílak az aktiváció és a gátlás szokásos értelmében szerepelnek.) A kismolekulán kívül egy másik szignál (g) jelenléte vagy hiánya szabályozza, hogy bekapcsoljon-e a transzport és a felhasználás, így a rendszer akkor is kikapcsolva maradhat, ha az adott kismolekula (cukor) megjelenik az extracelluláris térben, mert a sejt valami jobban hasznosuló másik molekulával kiváltja[6].

Ezekben a rendszerekben a cukor transzporter (pl. LacY) és a lebontó enzim (pl. LacZ) génjei egyidejűleg szabályozódnak. Azonban számos cukor, így a D-galaktóz hasznosítási hálózat (*gal* rendszer) esetében is sokkal nagyobb komplexitást tapasztalunk, például több transzporter, több regulátor, illetve számos lebontó, módosító enzim vesz részt a cukor hasznosításában[9]. Ebben a hálózatban a gének különböző operonokba rendeződnek, ami lehetővé teszi, hogy ezek különböző módon legyenek szabályozva. Sőt, előfordul az is, hogy egy operonon belül, annak génjei különböző módon szabályozódnak (ezt nevezik szuboperonikus szabályozásnak)[10].

2.2.1 A szabályozó gének

A rendszer bemutatását a szabályozó génekkel kezdem, mert így a későbbi elemek leírásánál hivatkozni tudok az itt leírtakra.

A galaktóz rendszerben két, egymáshoz szekvencia szinten és funkcióban is nagyon hasonló represszor, a gal represszor (GalR) és a gal izorepresszor (GalS) tölti be a szabályozó gén szerepét, melyek a LacI-t is magában foglaló regulátorcsaládba tartoznak. Ezek a fehérjék helix-turn-helix DNS-kötő motívummal és galaktózkötő doménnel rendelkeznek, és galaktózhoz nem kötött formában dimerként kötődnek (2.4 ábra.) az operátorhelyükhöz. A galaktóz kötődése által bekövetkező konformációváltozás hatására a két represszor DNS-affinitása lecsökken, így disszociálnak operátorukról. A GalR és a GalS nagyfokú homológiát mutat, és operátor specifitásuk is hasonló (ezért az izo az elnevezésében)[9, 11-13].

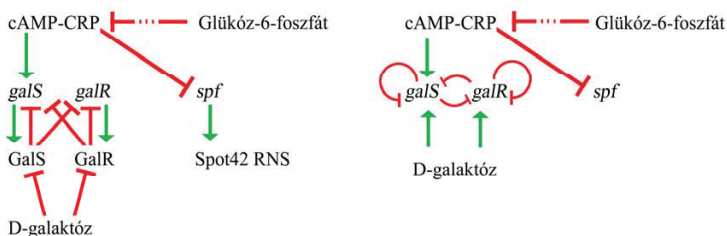


2.4 ábra. DNS-hez kötött GalR dimer modellje. A DNS az ábra felső részén szürke hélix. A kék és lila színnel jelölt monomerek azonosak. A pirossal jelölt felszínen találhatóak a DNS hurok kialakulásához szükséges tetramerizációban részt vevő aminosavak[14].

A GalS represszorral ellentétben a GalR fehérje képes tetramerizálódni is[15]. A GalR represszor a transzkripciót többféle mechanizmussal is képes gátolni (térbeli gátlás[11], kontakt gátlás[16-18], DNS hurok képződés[19-22]), valamint bizonyos körülmények között aktivátorként is működhet[16, 17, 23].

A szabályozásban részt vesz a cAMP is, a sejt számos alrendszerének közös szignálmolekulája, ami a sejt szénforrás hiányos, illetve magas szénforrás igényű állapotát közvetíti. A glükóz a preferált szénhidrát, ennek hiányában a sejt számos rendszerben, így a galaktóz hálózatban is változások állnak be. A cAMP az azt érzékelő CRP fehérjéhez (cAMP Receptor Protein) kötődve, cAMP-CRP komplex formájában fejt ki hatását. A CRP fehérje a cAMP-kötött formájában ismeri fel a DNS-en a kötőhelyét, és transzkripció regulátorként működik. Aktivátorként és represszorként is működhet[24].

A rendszer finomhangolásában egy kis RNS is részt vesz, a Spot42 RNS, mely az *spf* gén terméke. Ezt a cAMP-CRP szabályozza negatívan, azaz az *spf* gén a sejt glükózban jól ellátott állapotában aktiválódik[25]. A 2.5 ábra a *gal* rendszer szabályozását mutatja be. A rendszer ismertetése során folyamatosan bővitem majd az ábrát, míg eljutok a ma ismert teljes képig.



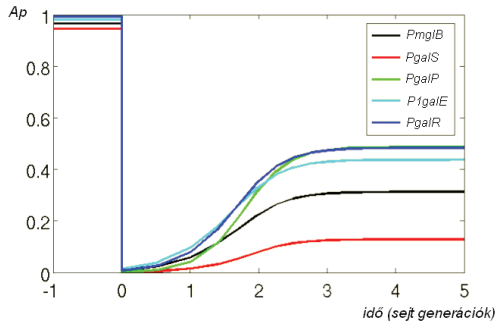
2.5 a) és b) ábra: A gal rendszert szabályozó elemek, és azok egymás közötti kölcsönhatásai. Az a) ábrán a gén-géntermék logikai kapcsolatot az aktiválás szimbólumával jelöljük. Az ábrán nem csak gének és géntermékek, hanem szignálmolekulák is láthatóak, amik a géntermékeken keresztül hathatnak a génekre. A glükóz-6-foszfát nem közvetlenül hat a cAMP-CRP komplexre, ezért a szaggatott vonalas ábrázolás. A b) ábra összehasonlításként azt mutatja meg, hogy milyen lenne mindez abban a szokványos, egyszerűsített jelölési rendszerben, amit általában használnak genetikai kölcsönhatások leírásában, melyben közvetlenül a génekre írják fel az egymással való kölcsönhatásokat. A dolgozat során ezen túl a csak az a) ábrának megfelelő ábrázolást fogom használni.

Korábbi munkánk során foglalkoztunk azzal, hogy miért lehetett evolúciósan szükséges a gal rendszerben két ennyire hasonló represszor, mint a GalR és a GalS. Az már ismert volt, hogy a két represszor közötti legfőbb különbség az, hogy a GalS sokkal erősebben represszálja saját magát, mint a célgénjeit (ezt nevezzük domináns negatív autoregulációnak)[26].

In vitro mért adatokon végzett matematikai modellezési eljárás segítségével elemeztük általánosságban a nem domináns-negatív és a domináns negatív önszabályozású transzkripciós faktorok (TF-ok) viselkedését különböző biokémiai paraméterek, mint a TF és a különböző operátorai közti disszociációs állandók, vagy a sejtbeli maximális TF koncentráció változtatásával. Ez a modell azt mutatta, hogy egy általános TF-célgén rendszer viselkedése steady-state helyzetben két dologtól függ:

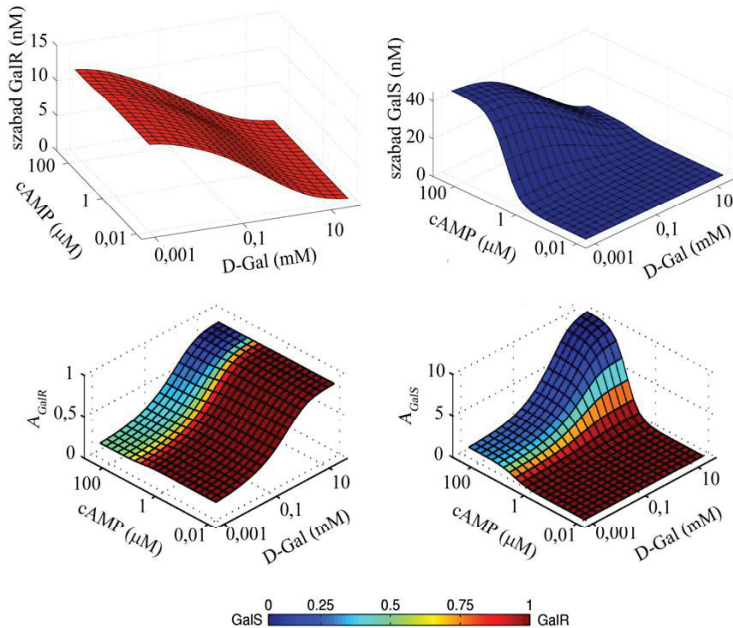
- Mennyire domináns (vagy nem az) a TF negatív önszabályozása a célgénhez képest, azaz a TF-nak a célgén operátorán mért disszociációs állandója (K) hogyan aránylik a TF-nak a saját magát szabályzó operátorán mért disszociációs állandójához (K_{self}), hiszen a domináns kifejezés éppen azt jelenti, hogy K_{self} kisebb, mint K .
- Mennyire magas maximális TF koncentráció lehetséges a sejtben, a K_{self} értékének arányában.

A modell megmutatta, hogy steady-state helyzetben a domináns negatív autoregulált TF hatékonysága elmarad a GalR-hez hasonló „hagyományos” TF-ok (melyek autoregulációja nem domináns) hatékonyságától, mivel az inducer cukor hiányában a TF domináns negatív autoregulációja egy olyan szint alatt tartja önmaga átírását, így a sejtbeli TF-ok koncentrációját, ami még nem elégséges a célgének represszálásához. Ugyanakkor egy dinamikus helyzetben, amilyen egy hirtelen történő cukormegvonás esetén (2.6 ábra) a sejt gyorsabban, hatékonyabban reagál ha egy ilyen represszor is jelen van, mivel az azt megelőző, cukorban gazdag periódus során nagy mennyiségű TF halmozódik fel a sejtben, ami a hirtelen cukormegvonás után átmenetileg gyorsan és erősen képes represszálni a célgénjeit. A GalS evolúciós haszna tehát valószínűleg abban áll, hogy segítségével a sejt sokkal gyorsabban tudja kezelni a hirtelen bekövetkező változásokat[26].



2.6 ábra. A GalS válasza a hirtelen galaktóz elvonásra, szimuláció. Az ábra GalS célgénjeinek szimulált promóter aktivitását mutatja be az idő függvényében. T=0 időpontban a galaktóz szint 10 mM-ről 0-ra esik. A modell szerint a korábban felhalmozódott nagy mennyiségű inaktív GalS aktívvá válik a galaktóz megvonás hatására. A hatás átmeneti, kb. 2 sejt generáción keresztül figyelhető meg[26].

Krishna és mtsai 2009-ben közöltek egy *in vitro* mérésekre épülő modellt, ami a cAMP és a D-galaktóz (a rendszer két bemenő jele) függvényében modellezi egy részről a rendszer promótereinek aktivitását, más részről a szabad represszorok mennyiségét a sejtben[27]. A 2.7 ábrán ennek a munkának a szabad (cukorhoz és operátorhoz sem kötött) represszorok intracelluláris mennyiségére vonatkozó modellje látható. A szabad represszorok mennyiségének meghatározása – mint később látni fogjuk – a rendszer viselkedését leíró modell kulcsfontosságú része.



2.7 ábra. A sejten belüli szabad (galaktózhoz és DNS-hez nem kötött) GalR és GalS mennyisége (fent), valamint a két represszorgén promóteraktivása (lent) a cAMP és a D-galaktóz függvényében. A promóteraktivitás mértéke (a lenti ábrákon a függőleges tengely) a P_{galR} promóter maximumához (=1) viszonyítva van ábrázolva. A promóteraktivást bemutató „lepedőábra” szinkódolása azt mutatja, hogy az adott operátoron az adott állapotban melyik represszor köt jellemzőbben. Így a kék pozíció azt jelenti, hogy az adott pozícióban (tehát az adott cAMP és D-galaktóz szintek esetén) és amennyiben az adott operátoron kötött represszor van, akkor ez az esetek 100%-ában a GalS, míg például a zöld pozíciókban az operátorhoz kötött represszor, ha van ilyen, 50-50%-ban a GalR és a GalS[27].

2.2.2 A cAMP-CRP komplex

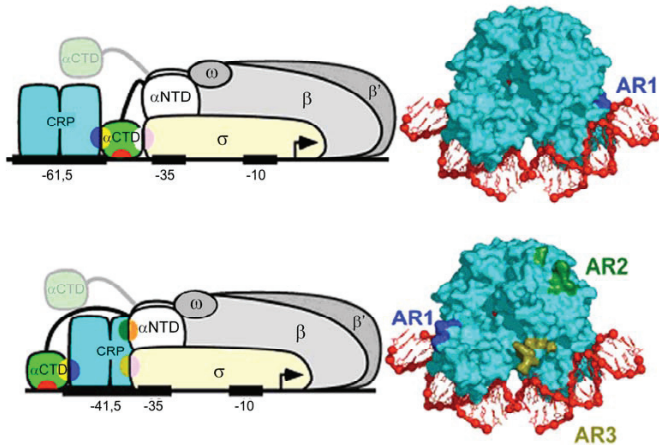
A két represszoron kívül a rendszert szabályozó fontos transzkripciósi faktor a cAMP Receptor Protein (CRP), más néven katabolit aktivátor protein (catabolite activator protein, CAP). Ez az *E. coli* egyik általános szabályozó fehérjéje, effektor a cAMP. A cAMP-kötött CRP (azaz cAMP-CRP) dimer formában köti a DNS-t helix-turn-helix DNS-kötő motívum segítségével, illetve meg is hajlítja azt. [28, 29] A CRP kötőhelyének konszenzus szekvenciája a DNS-en: TGTGAN₆TCACA. [30]

A cAMP-CRP képes aktiváló és gátló szerepet is betölteni. Az általa aktivált promótereket két csoportba osztják, a CRP-függő promóterek I. és II. osztályába (2.8 ábra). Az I. osztály – melyek közül az egyik legjobban leírt a *lac* promóter – esetén a CRP kötőhelye a promóter központi részéhez képest upstream helyezkedik el. (A CRP kötőhely közepe a -61,5 pozíció.) Ebben a promóter osztályban a CRP és az RNS-polimeráz (RNAP) közt egyetlen fehérje-fehérje kapcsolat jön létre, és egy szimpla toborzási mechanizmus által történik az aktiválás.

A II. osztályban, melybe a $P_{I_{galE}}$ is tartozik, a CRP kötőhelye átfed a promóter központi részének -35 régiójával, a $P_{I_{galE}}$ esetén a kötőhely középpontja a -41,5 pozíció. A II. osztályba tartozó promóterek esetén háromféle fehérje-fehérje kölcsönhatás jön létre a CRP és az RNAP között, melyek a toborzásban és a toborzás utáni lépésekben is segítik az RNAP-t, így hozzá járulnak a zárt komplex, majd abból a nyílt komplex létrejöttéhez. [28]

A cAMP-CRP komplex az aktiváláson kívül gátló funkciót is elláthat. Ennek két módját írták le eddig: a CRP-kötőhely pozíciója miatt a DNS-kötött cAMP-CRP komplex gátolja az RNAP kötődését a promóterhez, vagy egy másik aktivátor fehérjének a kötődését gátolja.

Az RNAP kötődésének gátlására több példát is ismerünk, ilyen a glükóz dehidrogenáz (*gcd*) promótere[31], vagy a *galETKM* operon $P_{2_{galE}}$ promótere[32]. Az aktivátorral való interferenciára példa a glutamát szintáz (*gluBDF*) operon gátlása, ahol a cAMP-CRP az IHF (Integration Host Factor) fehérjével interferál[33].



2.8 ábra. A cAMP-CRP és az RNAP kölcsönhatása az I. (fent) és II. (lent) osztályba tartozó promótereken. Jobb oldalon a cAMP-CRP dimer DNS-kötött állapotában látható, eltérő színekkel a cAMP-CRP és az RNAP közti kölcsönhatásban részt vevő aktivációs régiók (AR1, AR2 és AR3) [28]

2.2.3 Szabályozó régiók

A 2.9 ábrán a rendszer – alább ismertetett – génjeinek szabályozó régióit mutatom be egymáshoz illesztve. A szabályozó régiók közül jól felismerhető a GalR operátorhelye, mely általában a transzkripciós start helytől körülbelül 65 bp upstream helyezkedik el, kivéve a *galR* gént, melyben ez közvetlenül a starthelytől downstream helyezkedik el. Konszenzus szekvenciája GTGKAANCGNTTMCAC[13]. A cAMP-CRP kötőhely ezekben a génekben 40 bp upstream helyezkedik el a transzkripciós starthelytől, a *galR* gén kivételével, melyben nem található[11]. Konszenzus szekvenciája TGTGAN₆TCACA [34]. A -35 box-nak nevezett promóter elem kizárólag a *P_{galR}* promóterben található meg.

(ATP-binding cassette) transzporter családba tartozik. Ez a rendszer a galaktózon kívül glükózt is képes a sejtbe transzportálni. [9]

A kis affinitású ($K_m = 50\text{-}450 \mu\text{M}$), ATP-t nem igénylő (galaktóz/ H^+ -szimporter) transzportrendszer a *galP* gén terméke, a galaktóz permeáz (GalP), ami a *P_{galP}* promotérről íródik át[42]. A GalP a 15000-nél több tagot számláló MFS (Major Facilitator Superfamily) családba tartozik, ahova a laktóz permeáz (LacY) vagy a Glicerin-3-foszfát transzporter (GlpT) is, és a amely család általános jellemzője a kétszer hatos csoportba rendezett 12 transzmembrán α -hélix. Kristályszerkezete alapján a GalP funkcionális trimert alkot. [43]

További, nem specifikus módon is transzportálódhat D-galaktóz a sejtbe de ezekről az utakról jelen dolgozatban nem lesz szó, mivel egyrészt nem hatékonyak, másrészt pedig nem állnak galaktóz szabályozás alatt [44]. A két fő transzportrendszer szabályzó régiói eléggé hasonló struktúrájúak, azonban a szabályozásuk eltérő (2.9 és 2.10 ábra)[11].

```
galP gttagtGTGAAGCGATTACACTgaTGTGATTTGCTTCACAtcttttttaactgtaactcaccTATCTTaattcacAataaaaaataacca  
mgLB aggcgATGTAAACCGCTTTCAAtcTGTGAGTGTTCACAGtatcttaacaactgtgatagcTATGATtgcaccGtttaacgcttgytaac
```

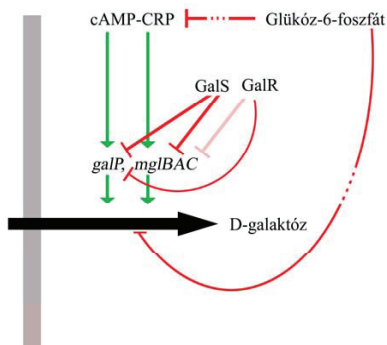
2.10 ábra. A 2.9 ábrának (a rendszer szabályzó régiói) a két transzportrendszerre vonatkozó része, ugyanazokkal a jelölésekkel. (Piros: GalR és GalS kötőhely, zöld: cAMP-CRP kötőhely, kiemelt: transzkripció start hely.)

A galaktóz transzportjára akkor van szükség, ha van elérhető extracelluláris D-galaktóz, valamint ha hiány van a D-galaktóznál preferáltabb szénforrásokból. Ennek megfelelően D-galaktóz hiányában mindkét transzportrendszert represszálja a GalR (és a GalS), valamint glükóz hiányában aktiválja a cAMP-CRP komplex. A cAMP-CRP általi aktiválás nélkül nem figyelhető meg alapaktivitás, ellenben a nagy affinitású *mgLBAC* operon esetén a GalR általi represszió gyengébb, így D-galaktóz hiányában is aktiválja a cAMP-CRP komplex. Ez biztosítja azt, hogy glükóz hiányában (azaz a cAMP-CRP általi aktiválás esetén) a sejt képes legyen extracelluláris D-glükózt és D-galaktózt is importálni, ha akár csak μM -os koncentrációban is jelen van a környezetben. Az *mgl* rendszer expressziója nélkül a sejt sokkal lassabban érzékelné, hogy D-galaktózban gazdag környezetbe került. Amennyiben a transzportált D-galaktóz elegendő ahhoz, hogy a represszorok inaktíválódnak, akkor a GalP átveszi a galaktóz transzportálás feladatát[11].

A környezetben elérhető galaktóz mennyisége erősen befolyásolja a cAMP-CRP intracelluláris szintjét. Kontrollált koncentrációjú galaktóz-limitált tápoldatban (chemostat)

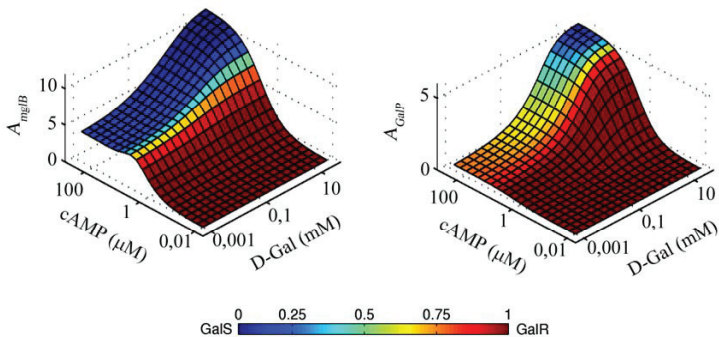
növesztett sejtekben jelentősen magasabb cAMP szintet figyeltek meg, mint magas galaktóz tartalmú kultúrában (batch) növesztett sejteknél[45]. Egyelőre azonban nem ismert, hogy a nem PTS cukrok hasznosítása milyen mechanizmussal a sejt cAMP szintjét befolyásolni[46, 47].

A galaktóz transzportot a glükóz-6-foszfát jelenléte (illetve valójában a nagymértékű képződése) egy a cAMP-CRP rendszertől független módon is gátolja, az inducer exclusion (inducer kizárás) nevű folyamat során. Ez a rendszer a tápoldatban lévő egyes nem preferált cukrok transzportereit gátolja, azaz a cukrokat (az inducer-eket) kizárja a sejtől. A folyamat a foszfoenol-piruvát-szénhidrát foszfortranszferáz rendszer (PTS) Enzyme II (EII) fehérjéjének az A doménje (EIIA^{Gk}) segítségével jön létre, mely defoszforilált állapotban képes megkötni és blokkolni egyes cukrok (így a laktóz és a galaktóz) transzporterét. A PTS rendszer a glükóz foszforilálással összekapcsolt transzportálását végzi. Az EIIA^{Gk} foszforilálódása egy foszfátcsoport transzfer része, mely során egy foszfátcsoport a foszfoenol-piruvátról több fehérjekomponensen át a sejtbe bekerülő glükózra kerül (a közttes lépések reverzibilis folyamatok). Ilyen módon az EIIA^{Gk} foszforiláltsága (és ezzel együtt az inducer exclusion mértéke) a PTS rendszer mindenkor aktívításának mértékétől függ. [48] A 2.11 ábrán a gal rendszer transzport alrendszerének szabályozását mutatom be.



2.11 ábra A gal rendszer transzport operonjainak szabályozási kapcsolatai. A nagy affinitású *mglBAC* operont a GalR gyengébben represszálja, így a környezetben levő D-galaktóz bejuthat a sejtbe. cAMP-CRP általi aktiválás hiányában egyik transzportrendszernek sincs aktivitása. A glükóz-6-foszfát a cAMP-CRP rendszeren kívül inducer exclusion útján is gátolja a galaktóz transzportot, ami az előzőhöz hasonlóan közvetve történik, ezt is szaggatott vonallal jelzem.

A már említett, Krishna és mtsai által közölt 2009-es publikációból[27] a két transzportrendszer promóter aktivitása a 2.12 ábrán látszik. Az ábrán is jól kivehető, hogy az *mgIBAC* operonnak – ellentétben a *galP* génnel – részlegesen aktív galaktóz hiányában, de magas cAMP szintnél.



2.12 ábra. A kis és nagy affinitású transzportrendszer promóter aktivitása a cAMP és a galaktóz függvényében. A promóter aktivitás mértéke (Z tengely) a P_{galR} promóter maximumához viszonyítva van ábrázolva (ld. 2.7 ábra). A promóter aktivitást bemutató „lepedőábra” szinkódolása azt mutatja, hogy az adott operátoron az adott állapotban melyik represszor köt jellemzőbben, így a kék pozíció azt jelenti, hogy az adott pozícióban, tehát az adott cAMP és D-galaktóz szintek esetén, és amennyiben az adott operátoron kötött represszor van, akkor ez az esetek 100%-ában a GalS, míg például a zöld pozíciókban az operátorhoz kötött represszor, ha van ilyen, 50-50%-ban a GalR és a GalS[27].

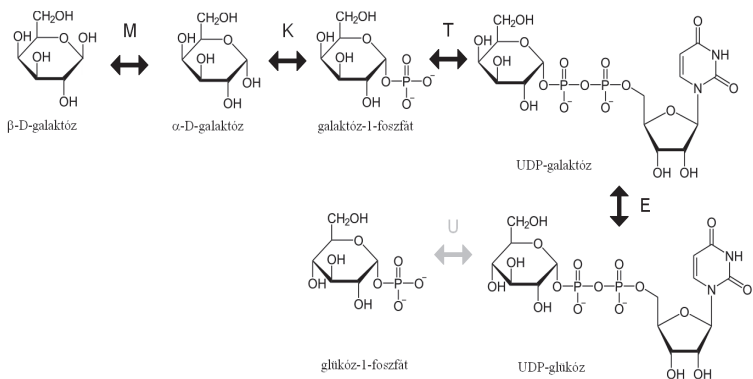
2.2.4 Metabolizmus gének és szabályozásuk

Az *E. coliban* a D-galaktóz hasznosítása amfibolikus módon történik (Leloir útvonal), mivel a D-galaktózt a sejt energiaforrásként (katabolikus út) és a sejtfal építőelemeként (anabolikus út) is használja.

A galaktóz metabolizmusért felelős öt eddig azonosított gén közül négy (*galE*, *galT*, *galK*, *galM*) egy operonba, a *galETKM* operonba (vagy *gal* operonba) szerveződik ebben a sorrendben[49]. Az ötödik gén, a *galU*, önálló operonként a kromoszóma egy másik pontján helyezkedik el[50], és nem áll galaktóz szabályozás alatt.

A *galM* gén a galaktóz mutarotáz enzimet kódolja[49], ami a D-galaktóz α és β anomerének átmenetét katalizálja a megfelelő α anomert szolgáltatva a következő lépéshez (az egyes katalitikus lépések a 2.13 ábrán láthatóak). A *galK* gén a galaktokináz enzimet kódolja, ami az α -D-galaktóz foszforilálásával galaktóz-1-foszfátot hoz létre[51]. A *galT* gén a galaktóz-1-foszfát uridililtranszferáz enzimet kódolja, ez az enzim UDP-galaktózzá alakítja a galaktóz-1-foszfátot[52]. Az operon negyedik génje, a *galE* az UDP-galaktóz-4-epimeráz enzimet kódolja, ami az UDP-galaktóz és UDP-glükóz közti átmenetet katalizálja, ez az egyetlen gén a *galETKM* operonban, ami a Leloire-útvonal anabolikus részéhez tartozik[53].

A *galU* gén szintén egy uridililtranszferázt kódol, mégpedig a glükóz-1-foszfát uridililtranszferázt, ami a glükóz-1-foszfát és az UDP-glükóz közti átmenetet katalizálja, és ez utóbbin keresztül kapcsolódik a galaktóz rendszerhez[50].



2.13 ábra. A galaktóz felhasználása és kapcsolata a glükóz felhasználással a *galETKM* és a *galU* operonok által. Az enzimek nevére csak egybetűs rövidítéssel utalok. Bővebb magyarázat a szövegben.

A *galETKM* operon két átfedő promóterrel rendelkezik, melyek eltérően működnek: ezek a $P1_{galE}$ és a $P2_{galE}$ [19, 54]. A 2.14 a) ábrán (mely a 2.9 ábrának, az ide vágó részét tartalmazza) egymás alatt szerepel a két promóter, ezek valójában 5 bp elcsúszással egymással átfednek. A 2.14 b) ábráján a valóságnak megfelelően látható a két promóter illetve szabályozó régióik. Az operon szabályozásában részt vesz egy második GalR

operátorhely is, mely az a) és b) ábrákon nem látható, mert kívül esik az ábrázolt szekvencián. Ez az operátor az operon első génjében, a *galE*-ben helyezkedik el, ezért belső operátornak (O_I), míg a másikat külső operátornak (O_E) nevezzük. Az O_I felfedezése volt az első eset ami rámutatott hogy egy operátor egy struktúrgénen belül is előfordulhat, elvezetve a DNS hurok képződés felfedezéséhez[20].

$P1_{galE}$: tcttGTGTAACGATTCACtaaTTTATTCATGTCACActtttcgcatcttTATGcTATGGTtatttcAataccataa
 $P2_{galE}$: taaattcttGTGTAACGATTCACtaaTTTATTCATGTCACActtttcgcatcTTTGtTATGCTatggttAtttcatac

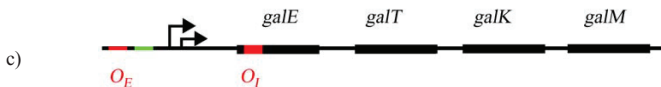
a)

ext. -10 -10 box tsp

tcttGTGTAACGATTCACtaaTTTATTCATGTCACActtttcgcatcTTTGtTATGCTATGGTtAtttcAataccataa

b)

1: ext. -10 -10 box tsp
 2: ext. -10 -10 box tsp

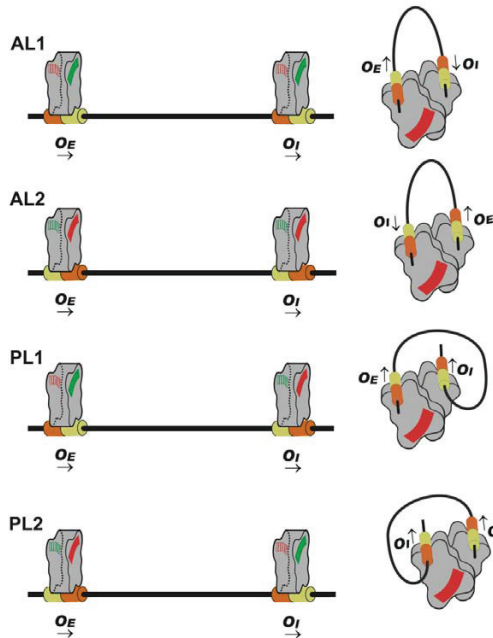


2.14 ábra. A D-galaktóz metabolizmus *galETKM* operonjának szabályozó elemei. Az a) ábrán a $P1_{galE}$ és $P2_{galE}$ promóterek a 2.9 ábrához hasonlóan külön szerepelnek, mégpedig a -10 box alapján illetve (a $P1_{galE}$ extended -10 box-a lényegében a $P2_{galE}$ -10 box-a). A *tsp* a transzkripció startpontot jelöli. A b) ábra szemlélteti a valós helyzetet, mégpedig hogy a két promóter 5 bp eltolással átfed, a szabályozó fehérjék kötőhelyei azonosak. A c) ábra sematikusan mutatja a b) ábrán bemutatott részeket, megtoldva az operon négy génjével. Így jól látható, hogy a *galE* génben egy második GalR-operátor (O_I) is található.

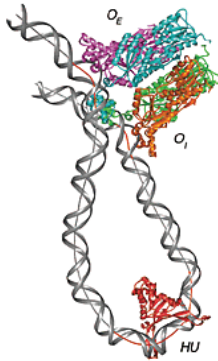
Az elemek számából is látszik, hogy a *galETKM* operon szabályozása a hálózat más tagjaihoz képest igen bonyolult. Előljáróban igen fontos tudni, hogy a két promóterről történő átírás más-más eredményhez vezet. Ha a $P1_{galE}$ promóterről történik a transzkripció, akkor a négy enzim egyenlő arányban képződik, és ez egyensúlyt teremt az anabolikus és a katabolikus utak között; míg a $P2_{galE}$ promóterről történő átírás esetén a GalE enzim nagyobb arányban képződik, ami a sejt működését inkább az anabolikus irányba tolja [55-58]. Ennek kivitelezése a következő módon történik.

D-galaktóz hiányában a szabad GalR dimerek mind a külső, mind a belső operátorhoz kötnek, és a HU (Heat Unstable Nucleoid Protein) nevű fehérje közreműködésével

DNS-hurkot képeznek és tetramer alakban stabilizálódnak (van egy HU kötőhely a DNS-en az O_E és az O_I között körülbelül félfúton, a transzkripció starthelyétől 6 bázisnyira downstream)[59, 60]. Ezáltal a két promóter a hurkok belsejébe kerül, és inaktívvá válik[19]. Mivel a GalR dimerek szimmetrikusak, így elméletileg négy lehetséges hurkok szerkezet alakulhat ki (2 paralell és 2 antiparalell), azonban ezek közül csak az egyik antiparalell szerkezet képes a transzkripciót hatékonyan gátolni (2.15 és 2.16 ábrák.)[21, 22, 61, 62].



2.15 ábra. A Gal represszósóma létrejöttének lehetőségei és hatásuk a transzkripció gátlására. A bal oldali rész a GalR heterodimerek négy lehetséges elrendeződését mutatja egymáshoz képest. A represszor fehérjén (szürke alakzat) a zöld csík a funkcionáló, a piros csík a nem funkcionáló interakciós felszín jelképezi, vagyis a két represszordimer a zöld felszíneken keresztül tud tetramerizálódni. A nyilak az operátor orientációját mutatják, az operátorban a GTGTAANC szekvenciájú oldal narancssárga, a GTGGTANC oldal zöldes. Az ábra középső részén a kialakult loop-ok (paralell: PL1 és PL2; antiparalell: AL1 és AL2) felépítését mutatja.[21]



2.16 ábra. GalR tetramer által létrehozott antiparalell DNS-hurok, melyet az HU fehérje stabilizál [61].

Azonban ha nem jön létre a hurok, akkor az O_E -kötött GalR képes arra, hogy a $P2_{galE}$ promótert aktiválja, vagyis ebben az esetben a fehérje valójában aktivátorként működik (A PI_{galE} promótert továbbra is represszálja). Ennek módja hasonlít ahhoz, ahogy a fentebb bemutatott CRP működik. A GalR köti az RNAP α -alegységének C-terminális doménjét (α CTD), és ez a kötés az operátorhoz közelebb eső $P2_{galE}$ esetében aktivációt eredményez, míg az 5 bp-ral messzebb elhelyezkedő PI_{galE} gátláshoz vezet, mivel az RNAP-promóter nyílt komplexének sebességhatározó lépésénél termodinamikai akadályt képez. [16-19].

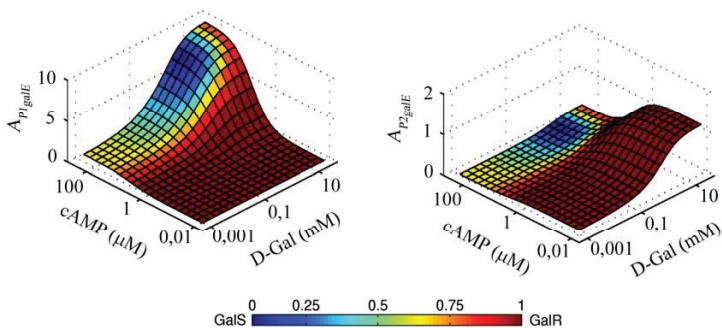
A GalR a külső operátort (O_E) erősebben köti, mint a belsőt (O_I), ami a $P2_{galE}$ promóternek kedvez[55].

A GalR tetramerizációja kooperatív módon jön létre. Mind a tetramerizációt, mind a GalR operátorhoz való kötődését befolyásolja a galaktóz jelenléte. Galaktóz teljes hiányában létrejön a DNS-hurokból, az azt összetartó GalR fehérjékből és a stabilizáló HU fehérjéből álló represszoszóma[14, 15, 21, 59, 61]. Kis mennyiségű galaktóz jelenlétében a tetramer kevésbé stabil, így a hurok stabilitása is csökken, ami kis mértékben lehetővé teszi a $galETKM$ operon átírását, elsősorban az $P2_{galE}$ promóterről[55].

A cAMP-CRP komplex szintén befolyással van az operon szabályozására, ez ugyanis a PI_{galE} -t aktiválja (de csak ha a GalR által létrehozott hurok nincs jelen), és a $P2_{galE}$ -t

gátolja[32]. Vagyis glükóz ellátottság esetén, még ha van is egy kis D-galaktóz a sejtben, a sejt inkább az anabolikus útvonalat választja a $P2_{galE}$ promóter által, és a D-galaktóz felhasználását elakasztva inkább az UDP-glükóz UDP-galaktózzá alakítását katalizálja. Ha van D-galaktóz, de glükózból hiány van, akkor a sejt áttér az energiatermelő felhasználásra a PI_{galE} által.

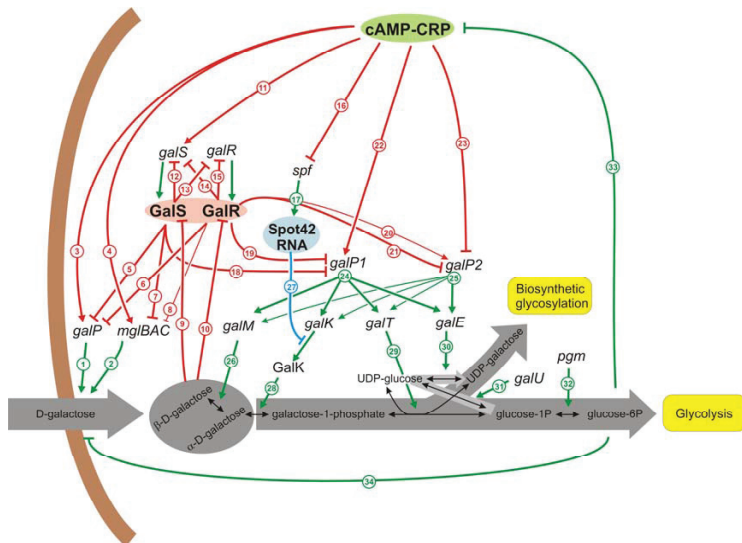
Ebbe a szabályozási rendszerbe a GalS nem nagyon szól bele, csak a PI_{galE} aktivitását represszálja részlegesen. A promóter aktivitásokra adott modellből (2.17 ábra) PI_{galE} -re vonatkozó (bal oldali) is jól látszik, hogy a GalS-nek csak nagyon korlátozott területen van tényleges represszáló hatása erre a promóterre, alacsony D-galaktóz és magas cAMP szint mellett a türkizkék és zöld mezők jelentik azt, hogy azon körülmények közt a represszió fele a GalS-től származik[27]. A 2.17 ábráról leolvasható, hogy a PI_{galE} magas cAMP-CRP és magas D-galaktóz szint mellett a legaktívabb, míg a $P2_{galE}$ alacsony cAMP-CRP szint és magas D-galaktóz szint mellett aktív. A modellből az is kitűnik, hogy a $P2_{galE}$ maximális aktivitása sokkal kisebb, mint a PI_{galE} -é, de ez önmagában semmit nem mond arról, hogy ténylegesen milyen arányban vannak a sejtben az operon enzimeit.



2.17 ábra. A *galETKM* operon két promóterének aktivitása a cAMP és a galaktóz függvényében. A promóter aktivitás mértéke a P_{galR} promóter maximumához viszonyítva van ábrázolva (ld. 2.7 ábra). A PI_{galE} esetén a szinkódolás azt mutatja, hogy az adott operátoron az adott állapotban melyik represszor köt jellemzőbben. A $P2_{galE}$ esetében azonban a kék-zöld-sárga mezők esetén a represszióért a cAMP-CRP felel, a GalS hatása marginális[27].

A fent részletezett szabályozási módokon kívül a rendszert posztranzkripciósan finomhangolja a Spot42 kis RNS, ami a *galK* Shine-Delgarno régiójához kapcsolódva gátolja az operonról készült mRNS translációját erről a pontról, s ezáltal a lebontó útvonal rovasára előnyt biztosít a *galE* által képviselt anabolikus útnak[10, 25].

Bár a *galU* gént említettem, mint a hálózat részét, szabályozása más keretek között zajlik, és jelen dolgozatnak nem része. A hálózat teljes eddig ismert szabályozását, beleértve az eddig külön bemutatott részeket is a 2.18 ábra szemlélteti összefoglalóan.

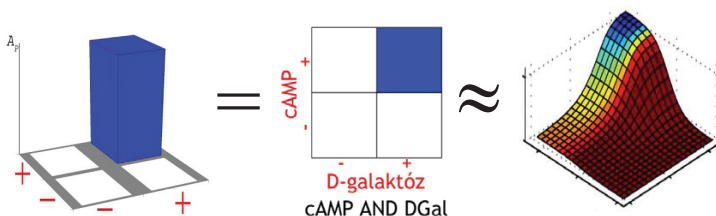


2.18 ábra. A gal rendszer szabályozási hálózata[11]. A részleteket ld. A szövegben.

2.3 A Boole-algebra, mint a szabályozás leírásának kvalitatív módja

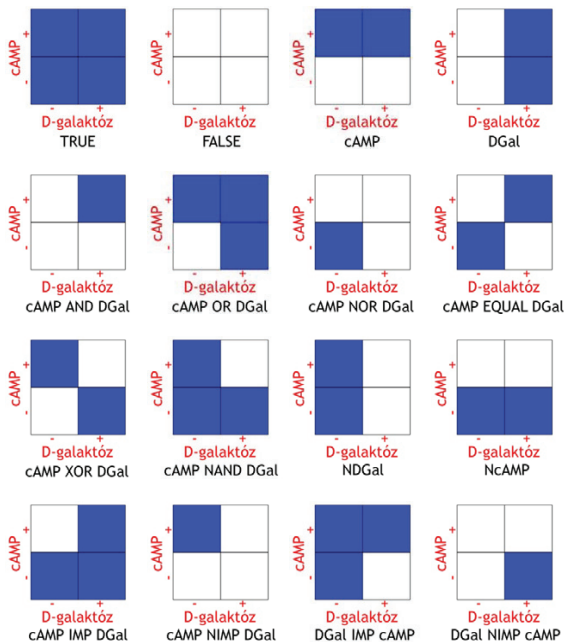
A gal rendszer tagjait – számos másik cukorhasznosítási hálózathoz hasonlóan – alapvetően két szignál szabályozza: a D-galaktóz és a cAMP. Mindegyik promóteren ez a két szignál összegződik valamilyen módon, vagyis minden egyes promóter esetében létezik egy rá jellemző mintázat, ami azt mutatja meg, hogy a két szignál hiányában vagy jelenlétében az adott promóter éppen aktív-e. Ha koncentráció átmenetek nélkül, azaz csak „van” és „nincs” állapotokat vehet fel a szignálok jelenléte, akkor 16 ilyen kombináció, azaz logikai kapu létezik, ezeknek a leírására alkalmasak az úgynevezett Boole-függvények[11, 63-68].

Ezeket ábrázolhatjuk egy háromtengelyű koordináta-rendszerben, amelyben az egyik tengelyen az egyik, a másik tengelyen a másik szignálmolekula hiánya vagy jelenléte szerepel, a harmadik tengelyen pedig megjelöljük, hogy az első két tengely függvényében hol tapasztalunk promóter aktivitást. A harmadik tengely valójában szinkódolást jelent, mellyel a koordináta-rendszer területén megjelöljük azokat a régiókat, ahol egy adott promóter aktív. Például ha egy promóter kizárólag akkor aktív, ha mind a cAMP, mind a D-galaktóz jelen van, akkor a logikai kapu ábrázolása a 2.19 ábrán látható módon fog kinézni.



2.19 ábra. A „cAMP AND DGal” logikai kapu azt jelenti, hogy az adott promóter kizárólag akkor aktív (a kék oszlop, illetve 2 dimenziósan ábrázolva a kék terület), ha mindkét szignál jelen van. A molekuláris mechanizmus, azaz hogy a promóter aktiválódása egy aktivátor fehérje miatt, vagy egy represszió feloldása miatt történik, illetve hogy az aktivitás hiánya egy represszor jelenlétéből vagy egy aktivátor hiányából fakad, az ebből az ábrázolásból nem derül ki. A harmadik ábrán egy már korábban látott promóter aktivációs ábrával (P_{galP}) szemléltetem, hogy egy ilyen egyszerűsített megközelítés valójában körülbelül milyen függvénynek felelhet meg.

Természetesen a szignál integráció 16 lehetséges módjába beletartoznak azok, amelyeknél a promóter egyáltalán nem reagál a szignálokra, ezeknél vagy minden állapotban van aktiváció, azaz a két szignál tekintetében konstitutív promóterek (TRUE), vagy pedig soha nincs aktiváció, azaz elromlott, vagy egy általunk nem vizsgált szignállal negatívan szabályozott promóterek (FALSE). Ide tartoznak továbbá azok a lehetőségek is, amikor a tényleges reakció csak az egyik szignál jelenlététől vagy hiányától függ (cAMP, NcAMP, Dgal, NDGal). A maradék 10 logikai kapu esetén mindkét szignál befolyásolja az eredményt. Az összes lehetséges kaput a 2.20 ábra mutatja be.



2.20 ábra. A 16-féle Boole-logikai kapu. A kék terület azt jelenti, hogy egy adott promóter mely szignál kombinációk esetén aktív.

A transzkripció szabályozás matematikai leírásának alapjai

Jelen dolgozat szoros együttműködésben készült a koppenhágai Niels Bohr Institute, Center for Models of Life kutatóival. A matematikai modellek kidolgozása Sandeep Krishna és Alexander Hunziker munkáját dicséri. A modellezés matematikai alapjai nélkülözhetetlenek a dolgozat megértéséhez. A modellezés oroszlánrészét Alex Hunziker végezte, ezért a munka ezen részének bemutatására az ő Sense & Sensitivity – Bacterial Response in Good and Bad Times című PhD dolgozatán alapul[69].

A génszabályozási hálózatok szimulációja általában közönséges differenciálegyenletek sorozatával történik. Ezek számos folyamat – beleértve a transzkripciót, a translációt, fehérje-fehérje interakciókat és transzportot – szimulációjára alkalmasak, ha a résztvevő részecskék száma elég nagy ahhoz, hogy a zaj ne játszhasson kritikus szerepet a rendszer dinamikus viselkedésében. Azokban az esetekben, melyekben bizonyos résztvevő molekulák száma nagyon alacsony (jelentősen kisebb 100-nál), valamilyen sztochasztikus megközelítésre van szükség, például a Gillespie algoritmust[70] használó szimulációra, esetünkben ilyenekre nem lesz szükség.

A transzkripciós faktorok által szabályozott gének transzkripciója dinamikus folyamat, melyben a transzkripciós faktorok újra és újra megkötik, illetve elhagyják az kötőhelyüket a DNS-en. Pozitív, azaz transzkripció aktiváció általi szabályozás esetén a promóter aktivitása arányos lesz azzal az idővel, amit a transzkripciós faktor aktív formája a DNS-hez kötve tölt (azaz a kötőhely *kötött hányada*). Negatív szabályozás esetén a számunkra releváns mennyiség a kötőhely *nem kötött, avagy szabad hányada*. Ezek a mennyiségek attól a disszociációs állandótól függenek, ami a transzkripciós faktor és a kötőhely közti komplex képződést írja le. Az egyenlet, ami a fehérje termelését és elbomlását írja le, így adódik:

$$\frac{dP}{dT} = A_p - \frac{1}{\tau} \cdot P \quad 2.i$$

Ez az egyenlet egy adott P fehérje koncentrációváltozását írja le az idő (T) függvényében A_p promóter aktivitás, τ átlagos fehérje élettartam és P aktuális fehérje koncentráció paraméterekkel. Minden pozitív kifejezés az egyenletben a fehérje termelését modellezi, a negatív kifejezés a bomlást jelenti. A fehérje termelése arányos a fehérjét kódoló gén promóter aktivitásával (A_p). Katalizált bomlás vagy a sejtből való fehérjeexport hiányában

a koncentráció kizárólag spontán bomlással vagy a sejtosztódás általi hígulással csökkenhet. A 2.i egyenlet csak a bomlást veszi figyelembe, mely arányos az aktuális fehérjekoncentrációval, és a bomlási rátával (ami τ , azaz az élettartam reciproka). Ez az egyenlet egyetlen kifejezésben írja le a teljes fehérje képződésének folyamatát a géntől a fehérjéig. Ennél komplexebb egyenletek is felírhatóak, melyek a transzkripciót és a translációt külön folyamatként írják le, vagy akár leírhatóak olyan alfolyamatok, mint az RNS polimeráz kötődése a promóterhez, vagy a tRNS toborzása a transláció során. Egyelőre egyszerűen elfogadjuk azt, hogy a gén bármilyen aktivitása végső soron a fehérje koncentrációjának ezzel arányos növekedéséhez fog vezetni, ami számos esetben elégséges, de szükség esetén kiegészíthető.

Több lehetőség is létezik, hogy a fent említett, és a promóter aktivitások szempontjából nagy jelentőségű értéket, a transzkripció faktor kötőhelyének a kötött illetve nem kötött hányadát kifejezzük. Ezek közül a munkánkhoz a tömeghatás törvényén alapuló megközelítést használtuk, ezért ennek bemutatására szorítkozom.

Vegyünk egy T transzkripció faktorot, és ennek operátorát a DNS-en, amit O -nak nevezünk el. (A matematikai levezetés nem változik, ha aktivátorról van szó, így a DNS kötőhelyet ezekben az esetekben is O -val jelöljük, mintegy kiterjesztve az operátor fogalmát bármilyen transzkripció faktor DNS-en levő kötőhelyére.) A kettő közti komplexképzés egyensúlyi folyamatként így írhatjuk le:



ahol az előre és a vissza irányuló reakciók rátáját az adott oldalon álló reagensek koncentrációjának, illetve a reakció megfelelő oldalához tartozó arányosságai együtthatójának a szorzata írja le. Egyensúlyi helyzetben a két ráta egyenlő, ami ezt adja:

$$k_{\text{on}}[T][O] = k_{\text{off}}[OT] \quad 2.\text{iii}$$

Mivel a K disszociációs konstanst úgy definiáljuk, hogy $K = \frac{k_{\text{off}}}{k_{\text{on}}}$, ki tudjuk fejezni a kötött T -t úgy hogy:

$$[OT] = [O] \frac{[T]}{[K]} \quad 2. \text{iv}$$

Feltesszük, hogy a transzkripció faktorok száma sokkal nagyobb, mint az operátor helyek száma a DNS-en, és így a szabad transzkripció faktor koncentrációja nem érzékeny arra, hogy abból mennyi kötődött a DNS-hez. Emiatt egyedül a T változót használjuk a transzkripció faktor koncentrációjának számításakor, eközben a szabad operátor helyek koncentrációját a teljes számuk függvényeként fogjuk kifejezni. Eddig az $[O]$ a szabad operátorhelyek koncentrációját jelölte, mivel ezek azok, amik elérhetőek a komplexképződéshez. Ha az operátorok teljes mennyiségét (O_{tot}) a szabad (O_{free}) és a kötött (OT) mennyiségek összegeként definiáljuk:

$$[O_{\text{tot}}] = [O_{\text{free}}] + [OT] \quad 2. \text{v}$$

Akkor megkapjuk ezt:

$$[OT] = ([O_{\text{tot}}] - [OT]) \frac{[T]}{K} \quad 2. \text{vi}$$

Végül megoldva a komplexre, megkapjuk ezt a kifejezést:

$$[OT] = [O_{\text{tot}}] \frac{\frac{[T]}{K}}{1 + \frac{[T]}{K}} \quad 2. \text{vii}$$

Vagy ha a szabad DNS-kötőhelyek számát akarjuk megkapni (ami az aktivátorral szabályozott gének esetében lehet szükséges), akkor egy hasonló egyenletet kapunk.

$$[O_{\text{free}}] = [O_{\text{tot}}] \frac{1}{1 + \frac{[T]}{K}} \quad 2. \text{viii}$$

Kombinálva a operátorhelyek transzkripció faktorhoz kötött és nem kötött hányadára talált kifejezéseket a 2.i egyenlettel, egy még teljesebb matematikai leírást kapunk a folyamatról. Például a transzkripció aktiválás esetében, ahol a promóter aktivitása arányos a kötött hányaddal, felírhatjuk ezt:

$$\frac{dP}{dT} = \frac{A_{\text{basal}} + A_{\text{max}} \frac{[T]}{K}}{1 + \frac{[T]}{K}} - \frac{1}{\tau} \cdot P \quad 2.\text{ix}$$

Az egyenlet itt egy kicsit kiegészített formában szerepel, számításba véve, hogy egy aktivátor által aktivált promóter akár az adott aktivátor hiányában is mutathat alapaktivitást (A_{basal}).

Ugyanezzel a megközelítéssel ennél összetettebb helyzetekre is felírhatunk egyenleteket. Például ha a transzkripció faktor a kötőhelyét multimer formában köti, a kötőhely kötött hányadát így lehet kifejezni:

$$[OT] = [O_{\text{tot}}] \frac{\left(\frac{[T]}{K}\right)^h}{1 + \left(\frac{[T]}{K}\right)^h} \quad 2.\text{x}$$

ahol a h a Hill koefficiens, ami a fehérje kooperativitást jelzi ($h = 2$ egy dimer esetén, $h = 4$ egy tetramer esetén, stb.)

3. Célkitűzés

Munkánk célja az volt, hogy az *E. coli* galaktóz hasznosítási hálózatát felhasználva választ kapjunk az alábbi kérdésekre:

- A feast-famine ciklus során időben hogyan alakul az egyes promóterek aktivitása?
- Ez milyen sejten belüli cAMP-CRP és D-galaktóz szintet feltételez?
- Vajon a sejt készül-e a „rosszabb időkre”, van-e valami előre meghatározott program a promóterek aktivitási időzítésében, vagy kizárólag valami külső, esetleg belső szignálra reagál?
- Van-e összefüggés a szabályozási logikák és az időzítés között?
- Milyen szabályozási logikákat lehet megvalósítani a rendszer elemeiből?
- Hogyan lehet megvalósítani egy szabályozási logikát?
- Mekkora flexibilitást biztosít a rendszernek a szabályozó elemek átrendeződése, megváltozása?

4. Anyagok és módszerek

Kísérletünk menete röviden összefoglalva a következő volt: A vizsgált (természetes és mesterséges) szabályozó szekvenciák felhasználásával plazmid konstrukciókat hoztunk létre. Ezekről PCR reakcióval amplifikáltuk a promóteket és egyéb szabályozó régiókat, és vad típusú, valamint adenilát-ciklázt szintetizálni nem képes ($\Delta cyaA$) *E. coli* kromoszómára rekombináltattuk a *gusA* (β -glükuronidáz) riportergén saját szabályozó régiójának helyére. A GusA a LacZ-hez hasonlóan egy cukorbontó enzim, melynek létezik az X-Gal-hoz hasonló kék-fehér szelekciót adó szubsztrátja (X-Gluc, 5-bromo-4-kloro-3-indolyl- β -D-glükuronsav), valamint enzimkinetikai mérést lehetővé tevő szubsztrátja (PNPG, para-nitrophenyl- β -D-galaktopiranozid). A $\Delta cyaA$ háttérű baktérium konstrukciókat arra használtuk, hogy a promóterek logikai viselkedését vizsgáljuk, mert így lehetőség volt a cAMP mennyiségét kívülről szabályozni, és kék-fehér tesztet végezni. A vad típusú *E. coli*-ra épülő konstrukciókon a feast-famine ciklust modelleztük, és különböző növekedési fázisból vett mintákon a GusA riportert enzim aktivitását megmérve és matematikai modellbe illesztve a vizsgált promóterek időbeli viselkedésére következtettünk.

4.1 Törzsek

Név	Genotípus jellemzés	Referencia
<i>E. coli</i> K-12 MG1655	<i>F- λ-ilvG- rfb-50 rph-1</i>	NCBI RefSeq NC_000913
<i>E. coli</i> CH1200	MG1655 $\Delta cyaA$	[14]
<i>E. coli</i> XL-1 Blue	<i>endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F' [::Tn10 proAB⁺ lac^R $\Delta(lacZ)M15]$ hsdR17($r_{K}^- m_{K}^+$)</i>	Stratagene

4.2 Táptalajok és kiegészítők

LB

- Pepton 10 g/l
- Élesztő kivonat 5 g/l
- NaCl 10 g/l
- A szilárd táptalajba 12 g/l bakteriológiai agart teszünk.
- Autoklávban sterilizálható

SOB

- Pepton 20 g/l
- Élesztő kivonat 5 g/l
- NaCl 0,5 g/l
- NaOH segítségével a pH = 7-re kell állítani.
- Autoklávban sterilizálható
- Felhasználáskor 5/1000 térfogat steril, 2M MgCl₂-dal egészítjük ki (≈ 10 mM MgCl₂)

SOC

- SOB + 0,02 M glükóz.
- A glükózt a többi oldat sterilizálása után, szűrővel sterilizált törzsoldatból adjuk hozzá.

2x M63 minimál („alap”)

- KH₂PO₄ 100 mM
- (NH₄)₂SO₄ 15 mM
- FeSO₄ 1,7 μM
- pH = 7,0- ra beállítani KOH segítségével
- Autoklávban sterilizálható.

M63 minimál (kiegészített)

- Nem specifikus kiegészítők:
 - 1x M63 minimál („alap”)
 - 2 mM MgSO₄
 - 0,1% Amicase (aminosav keverék)
 - 1% B1 vitamin
- Specifikus kiegészítők:
 - D-galaktóz 0,4%
 - Zeocin: 8 µg/ml (a szokásosnál kevesebb)

Táptalaj kiegészítők végkoncentrációi:

- Zeocin: 80 µg/ml
- Ampicillin: 80 µg/ml
- X-Gal: 0,2%
- X-Gluc: 50 µg/ml
- D-Galaktóz: 8 mM
- Arabinóz: 0,23%
- cAMP: 0,16 mM

4.3 Klónozás

Az általános klónozási eljárásokhoz XL-1 Blue klónozó törzset használtunk (Stratagene), melyből az Inoue-féle módszer alapján készítettünk szuperkompetens sejtet, az alábbi recept szerint.

A törzset LB lemezen egy telepre kentük, melyből másnap 10 ml folyékony LB tápoldatba oltottunk egy kolóniát. Ebből 6-8 órányi 37°C-on, 255 rpm rázatás mellett történt növesztés után háromszor 500 ml párhuzamosan indított, 10 mM MgCl₂ tartalmú SOB tápoldatba oltottunk 1, 2 és 4 ml-t. Ezeket 19-22 °C-on, 200 rpm rázatás mellett növesztettük egész éjszaka, és másnap addig, míg az egyik kultúra elérte az OD₆₀₀ = 0,55

értéket. A másik két párhuzamos kultúrát eldobtuk. (A háromszoros túlbiztosításra az esetleges túlnövekedés miatt volt szükség.)

Minden szükséges oldatot jégbe hűtve előkészítettünk. A felnövekedett baktérium kultúrával 10 perc jégen való inkubálás után kezdtünk dolgozni.

A kultúrát 3000 xg relatív centrifugális erővel, 4°C-on, 10 percig centrifugáltuk. A felülúszót alaposan leszárítottuk, majd 2 percere fejfel lefelé steril papírvattára állítottuk, hogy minden folyadék maradéktalanul távozzon. A csapadékot 8 ml jéghideg, PIPES-sel kiegészített Inoue reagensben vettük fel. (Ennek receptjét a protokoll alatt közlöm.) A fenti módon centrifugáltuk, dekantáltuk és papírvattán szárítottuk. A csapadékot 4 ml, PIPES-sel kiegészített Inoue reagensben vettük fel, és 225 µl DMSO-t adtunk hozzá. Jégen inkubáltuk 10 percig, majd 100 µl-es térfogatú adagokba szétmérve, -70 °C-on tároltuk.

Szokásos hősokk transzformálással transzformáltuk, arra ügyelve, hogy ezeknek a sejteknek 90 másodperces hősokk az optimális, e fölött jelentősen romlik a transzformálás hatékonysága. A regeneráláshoz folyékony SOC médiumot használtunk.

Inoue reagens:

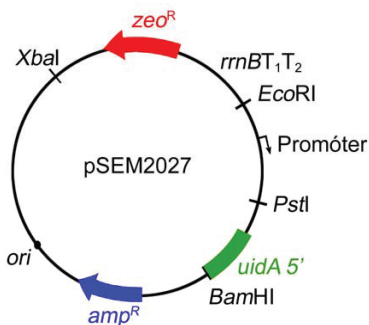
- KCl 250 mM
- CaCl₂ 15 mM
- MnCl₂ 55 mM
- Kuktában sterilizálás után -20°C-on tároltuk.
- Használat előtt 500 mM steril szűrt PIPES (pH = 6,7 KOH-oldattal beállítva) oldattal 1/100 térfogatnyival kell kiegészíteni. (PIPES végkoncentráció: 5 mM)

4.4 Promóter konstrukciók létrehozása

A szabályozó régió-promóter konstrukciókat szintetikus úton állítottuk elő egymással komplementer régiókat tartalmazó primer párokkal vagy az *E. coli* MG1655 törzséről történő PCR reakció segítségével. A primer párok esetében az egyik *EcoRI* (upstream), a másik pedig *PstI* (downstream) linker szekvenciát tartalmazott.

A szekvenciákat a Semsey Szabolcs által készített pSA850 alapú pSEM2027 vektoron hoztuk létre, melyet a 4.1 ábra mutat be. Erről az uidRzdn (5'-ACCCGGATCCTCA ATGCTGCCAGAGAGATTTTTTCAGAAAATGGATTTCACGGAATTCTCAGTCCTG CTCCTCGGCCAC-3') és a GusSeqDN (5'-TTCTTGTAACGCGCTTCCCAAC-3')

primerpárral amplifikáltuk a konstrukciókat. A termék a teljes zeocin-rezisztenciát adó *zeo^R* kazettát, az *rrnBT₁T₂* terminátor csoportot, a promotert és az *uidA* gén első 134 bp-nyi régióját tartalmazta. A zeocin-kazetta oldaláról a promotér segítségével az amplifikáció során építettünk a konstrukcióhoz egy 50 bp-nyi régiót, mely az *E. coli* MG1655 kromozómán az *uidA* gén vad típusú promotérétől upstream levő szekvenciával azonos. A későbbi rekombinációs lépéshez ez a kettő, legalább 50 bp-os szekvencia homológia szükséges.



4.1 ábra. A pSEM2027 vektor.

4.5 Riportertörzsek létrehozása

A riportertörzsek létrehozásához a λ -red rendszert használtuk[71]. A hőérzékeny replikációjú pKD46 plazmiddal transzformált és overnight növesztett baktériumot 0,23% arabinózzal indukálva OD600 = 0,5-ig növesztettük, majd elektrokompetens sejteket készítettünk belőlük (Ennek protokollját ld. alább.). A plazmidon a λ rekombinációs rendszer három génje, a *gam*, *bet*, *exo* található, melyek *P_{araBAD}* promotérről íródnak át.

Az 50 μ l-es adagokban tárolt elektrokompetens sejtekhez hozzáadtunk 5-7 μ l rekombinálandó PCR terméket, majd 5 perc jégen inkubálás után 1 mm-es réstávolságú küvettdában, a gyártó által ajánlott körülmények között (1,8 kV/mm; 1 sokk) végeztük az elektroporálást.

A baktériumokat másnap reggelig SOC médiumban rázás nélkül, szobahőmérsékleten inkubáltuk, majd másnap LB + 80 µg/ml zeocin lemezre szélesztettük és 37 °C-on inkubáltuk. A továbbiakban végig 37°C-on dolgoztunk a törzsekkel, hogy a pKD46 plazmid eliminációja megtörténhessen.

4.6 Elektrokompetens sejt készítése

LB tápoldatban overnight növesztett baktérium kultúrát 1:100 arányban, folyékony LB tápoldatba kihígítottunk, és 37°C-on (hőérzékeny törzsek esetén 30 °C-on), 250 rpm rázatás mellett, szükség esetén indukálva növesztettük OD₆₀₀ = 0,5-1,0 közötti értékre (0,6 az optimális). Az eljárás során minden összetevőt és sejtet jégen tartottunk, illetve jégbe hűtve előkészítettünk.

A felnőtt kultúrát centrifugáltuk: 4000 xg, 15 perc, 4 °C-on – minden centrifugálási lépés ezekkel a beállításokkal zajlott. A felülúszót alaposan eltávolítottuk, majd a csapadékot a kiindulási térfogattal egyenlő mennyiségű jéghideg, steril, 1 mM koncentrációjú HEPES oldatban vettük fel. (A HEPES autoklávban sterilizálható.) Újabb centrifugálás és a felülúszó eltávolítása után a kiindulási térfogat felének megfelelő 1 mM HEPES oldatban vettük fel, és újra centrifugáltuk. A csapadékot ezek után a kiindulási térfogat 1/50-ének megfelelő mennyiségű jéghideg, steril 10 %-os glicerinben vettük fel (a glicerin szintén autoklávozható), majd az utolsó centrifugálás után 1/100 térfogatnyi 10% glicerinben vettük fel a sejteket. Az ilyen módon betöményített és ionmentesített sejtszuszpenziót 50 µl-es térfogatú adagokban, -70°C-on tároltunk.

4.7 A szignál integráció vizsgálata *in vivo*

A létrehozott CH-1200 alapú riporter törzseken első lépésként vizsgáltuk a *gusA* gén működését. A vizsgált telepeket 50 µl LB-ben vettük fel, és ebből a híg szuszpenzióból 2-2 µl-t cseppentettünk az alábbi lemezekre, majd vizsgáltuk a kapott baktériumfolt színét:

- 1) LB + 80 µg/ml zeocin + 50 µg/ml X-Gluc
- 2) LB + 80 µg/ml zeocin + 50 µg/ml X-Gluc + 0,16 mM cAMP
- 3) LB + 80 µg/ml zeocin + 50 µg/ml X-Gluc + 8 mM galaktóz
- 4) LB + 80 µg/ml zeocin + 50 µg/ml X-Gluc + 8 mM galaktóz + 0,16 mM cAMP

4.8 Törzsek ellenőrzése szekvenálással

A megfelelő színreakciót mutató törzsek közül kromoszómális DNS-t tisztítottunk a Wizard Genomic DNA Purification Kit (Promega) segítségével, melyről a KpnT1T2 (5'-ATATATGGTACCAAGCTTCTGTTTTGGCGGATGAGA-3') és a GusSeqDN (5'-TTCTTGTAACGCGCTTCCCACCAAC-3') primerek felhasználásával amplifikáltuk az *uidA* rekombináns szabályozó régióját. A szekvencia meghatározását a Biomi Kft (Gödöllő) végezte.

4.9 β -glükuronidáz assay

Az overnight LB+zeocin tápoldatban növesztett baktériumtenyészetet 0,4 % galaktózzal és 8 $\mu\text{g/ml}$ zeocinnal is kiegészített M63 minimál („teljes”) tápoldatba oltottunk 1:1000 hígítással. Óránként mértük 600 nm-en az optikai denzitást (OD_{600}). Amikor az OD_{600} érték elérte a 0,2-t, minden mérés alkalmával kivettünk a kultúrából 1 ml-t, 10 percig 4°C-on 17000 xg relatív centrifugális erővel centrifugáltuk. A felülúszó eltávolítása után -70°C-ra helyeztük a mintákat.

Az enzimaktivitás méréshez 0,2 OD egységnyi mintát használtunk 500 μl térfogatban. Puffer gyanánt vitaminokat és aminosvakat nem tartalmazó 1X M63 minimál tápoldatot használtunk, melyet az esetleges fehérje transzláció gátlása érdekében 80 $\mu\text{g/ml}$ kloramfenikollal egészítettünk ki (ez a normálisan, szelekcióra használt mennyiség négyszerese). A baktériumokat 0,1 mg/ml lizozimmel kezeltük 15 percig jégen inkubálva, majd további feltárást végeztünk 250 μl feltáró oldattal szobahőmérsékleten. (Az oldatok receptje a protokoll alján található.) Az inkubációs idő lejártával a küvetákban előkészített és sötétben tartott 250 μl mérőoldathoz adtuk a feltárt szuszpenziót, és 15 percenként mértük a 405 nm-es optikai denzitást. A mérés alatt a mintákat tartalmazó küvetákat 37°C-on, 100 rpm rázatás mellett inkubáltuk. Ahol szükséges volt, a mérés időközzeit csökkentettük.

1X M63 minimál + Kloramfenikol

- KH_2PO_4 50 mM
- $(\text{NH}_4)_2\text{SO}_4$ 7,5 mM
- FeSO_4 0,85 μM
- pH = 7,0- ra beállítani KOH segítségével
- Kloramfenikol 80 $\mu\text{g}/\text{ml}$

Feltároló oldat

- Tris (pH = 8) 100 mM
 - NaH_2PO_4 12,8 mM
 - Na_2HPO_4 19,2 mM
- } = 32 mM foszfátpuffer, pH = 7
- DTT 8 mM
 - Polymyxin-B 0,2 mg/ml
 - CDTA 80 mM
 - Triton X-100 4 V/V %

Mérőoldat

- DTT 0,5 mM
 - EDTA (pH = 8) 1 mM
 - NaH_2PO_4 20 mM
 - Na_2HPO_4 30 mM
- } = 50 mM foszfátpuffer, pH = 7
- PNPG 125 μM

4.10 *In vitro* transzkripció

Az *in vitro* transzkripciót a Geanacopoulos és mtsi. által közölt módszer[72] alapján végeztük. Az *in vitro* transzkripcióhoz használt CRP fehérjét a Ryu és mtsai által leírt módon[73] tisztítottuk. Az RNS polimeráz (RNAP) σ^{38} alegységet a IMPACT rendszer (New England Biolabs) tisztítottuk, úgy, ahogy Shin és munkatársai megadják[74]. A tisztított σ^{38} alegységet 1:1 arányban kevertük az Epicentre-től vásárolt RNAP core enzimhez (C90500). A σ^{70} RNAP-t az USB-től vásároltuk.

A reakcióelegy összetétele:

20mM Tris acetát, pH 7,8

10 mM Magnézium-acetát

200 mM Kálium-glutamát

2nM DNS templát (supercoiled)

Ahol szükséges volt, a CRP-t 50 nM, a cAMP-t 100 μ M mennyiségben alkalmaztuk.

20 nM RNAP-t adtunk az elegyhez, majd 5 percig 37°C-on inkubáltuk. A transzkripció elkezdéséhez hozzáadtuk a nukleotid trifoszfátokat:

- 1 mM ATP
- 0,1 mM GTP
- 0,1 mM CTP
- 0,01 mM UTP és 33 mM [α - 32 P]UTP (3000 Ci/mmol; 5 μ Ci reakciónként)

A reakciót 10 perc inkubálás után állítottuk le az alábbi transzkripciós felvivő pufferrel:

- 0,025% Brómfenolkék
- 0,025% Xilencianol
- 0,01 M EDTA
- 90% Deionizált formamid

3 percnyi 90°C-os hőkezelés után 7%-os poliakrilamid-urea DNS szekvenáló géltre vittük a mintákat. Az RNS mintákat ImageQuantTM PhosphorImager (Molecular Dynamics CA) segítségével kvantifikáltuk.

4.11 A *gal* rendszer promóter aktivitásainak modellje

A modellezést Alex Hunziker és Sandeep Krishna végezte a 2009-ben közölt modell alapján [27]. A használt keretrendszerben a P fehérje koncentrációjának időbeli változása leírható az alábbi egyenlettel:

$$\frac{dP_{\text{tot}}}{dt} = v_p A_p(P_{\text{tot}}, g, c) - P_{\text{tot}}$$

P_{tot} a fehérje teljes mennyisége, v_p a fehérje maximális produkciós rátája, A_p a promóter aktivitása, g és c pedig a galaktóz és a cAMP koncentráció.

A fent említett modellből (ref[27]) explicit kifejezést kaphatunk az A_i aktivitásokra, ahol i a *gal* regulon bármely fehérréje lehet, s ezekkel a bevezetőben leírtak alapján tovább számolhatunk.

Az *mglB*, *galS* és *galP* gének esetében a promóter aktivitások a következő kifejezéssel írhatóak fel:

$$\frac{A_i}{A_{i, \max}} = \frac{1}{1 + \left(\frac{R}{K_{r,i}}\right)^{h_{r,i}} + \left(\frac{S}{K_{s,i}}\right)^{h_{s,i}}} \cdot \frac{\left(\frac{c}{K_{c,i}}\right)^{h_{c,i}}}{1 + \left(\frac{c}{K_{c,i}}\right)^{h_{c,i}}}$$

ahol $A_{i, \max}$ a maximális promóter aktivitás, R és S a szabad GalR és GalS mennyisége, K -val és h -val rendre a megfelelő (indexben jelölt) disszociációs állandót és hill-koeficienst jelöljük. Ez a kifejezés az adott promóter represszióját fejezi ki a két represszor által, valamint a cAMP-CRP általi aktivációt.

A P_{galR} promótert nem csak a szabad, hanem a galaktóz-kötött GalR (R_g) is képes represszálni, de a cAMP nem, ezért esetében a promóter aktivitásra ez az egyenlet írható fel:

$$\frac{A_{galR}}{A_{galR, \max}} = \frac{1}{1 + \left(\frac{R}{K_{r, galR}}\right)^{h_{r, galR}} + \left(\frac{R_g}{K'_{r, galR}}\right)^{h_{r, galR}} + \left(\frac{S}{K_{s, galR}}\right)^{h_{s, galR}}}$$

A *galETKM* operon $P_{I_{galE}}$ promóterére a következő két kifejezést használhatjuk, ami cAMP-CRP általi aktivációt, valamint a szabad és galaktózkötött GalR és a szabad GalS általi repressziót írja le:

$$\frac{A_{P_{I_{galE}}}}{A_{P_{I_{galE}}, \max}} = \frac{\left(\frac{c}{K_{c, OE}}\right)^{h_{c, OE}}}{1 + \left(\frac{c}{K_{c, OE}}\right)^{h_{c, OE}}} \cdot \frac{1 + 0,025 \left(\frac{R}{K_{r, OE}}\right)^{h_{r, OE}} + 0,025 \left(\frac{R_g}{K'_{r, OE}}\right)^{h_{r, OE}}}{1 + \left(\frac{R}{K_{r, OE}}\right)^{h_{r, OE}} + \left(\frac{R_g}{K'_{r, OE}}\right) + \left(\frac{R}{K_{r, loop}}\right)^{2h_{r, loop}} + \left(\frac{S}{K_{s, OE}}\right)^{h_{s, OE}}}$$

A következő egyenlet pedig a $P2_{galE}$ aktivitását írja le, az eddigiektől eltérően az alapaktivításra normálva a maximális aktivitás helyett. A promotert a GalS önmagában 1,2-szeres mértékben, a GalR 2-szeres mértékben aktiválja, a cAMP-CRP komplex represszálja. Ha a cAMP-CRP és a GalR együttesen kötődik, akkor az előbbi repressziója gyengébb.

$$\frac{AP2_{galE}}{AP2_{galE, basal}} = \frac{1}{1 + \left(\frac{c}{K_{c, P2galE}} \right)^{hc, P2galE}} \cdot \frac{1}{1 + \left(\frac{R}{K_{r, OE}} \right)^{hr, OE} + \left(\frac{R_g}{K'_{r, OE}} \right)^{hr, OE} + \left(\frac{R}{K_{r, loop}} \right)^{2hr, loop} + \left(\frac{S}{K_{s, OE}} \right)^{hs, OE} + \left(1 + 2 \left(\frac{R}{K_{r, OE}} \right)^{hr, OE} + 2 \left(\frac{R_g}{K'_{r, OE}} \right)^{hr, OE} + 1,2 \left(\frac{S}{K_{s, OE}} \right)^{hs, OE} + \left(\frac{R_g}{K'_{r, OE}} \right)^{hr, OE} \cdot \left(\frac{c}{K_{c, P2galE}} \right)^{hc, P2galE} + \left(\frac{R}{K_{r, OE}} \right)^{hr, OE} \cdot \left(\frac{c}{K_{c, P2galE}} \right)^{hc, P2galE}}}$$

Fenti egyenletekben egyaránt előfordul a GalR galaktózkötött és szabad állapota is. A kettejük közti átmenet a kettejük közti kötési állandó és Hill-koefficiens ismeretében így írható fel:

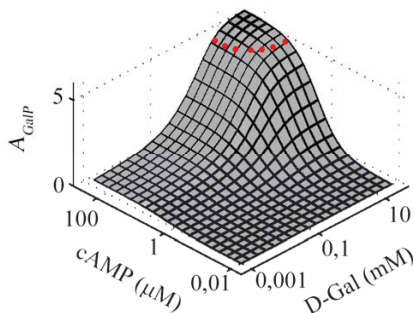
$$Rg = R \left(\frac{g}{Kr, g} \right)^{h, rg}$$

A represszor négy formája (szabad: R , galaktóz-kötött: R_g , operátor-kötött: R_{bound} valamint galaktóz és operátorkötött: $R_{g bound}$) adja ki a teljes GalR mennyiséget:

$$R_{tot} = R + R_g + R_{bound} + R_{g bound}$$

4.12 A D-galaktóz és a cAMP-CRP szintek meghatározása

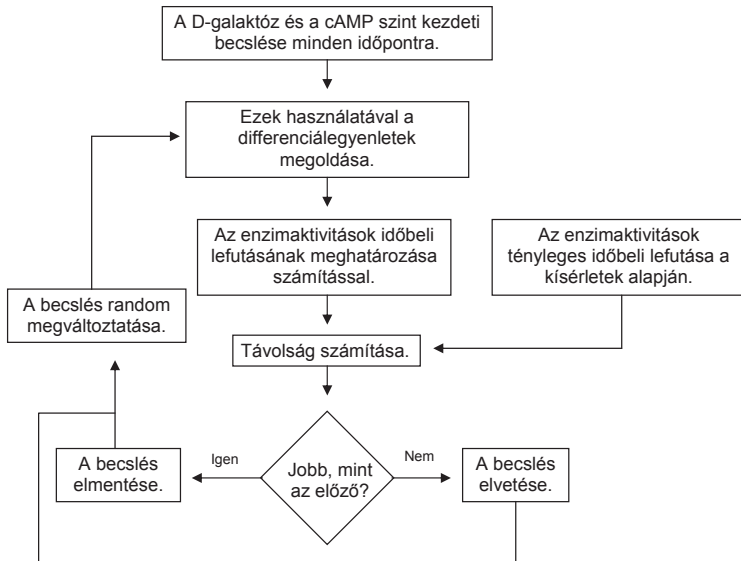
A sejtbeli D-galaktóz és cAMP-CRP szinteket a kísérletes eredményekből és a fenti modell alapján Alexander Hunziker a simulated annealing módszerrel határozta meg[69]. A probléma, amit a módszerrel meg kell oldani, hogy egy promóter aktivitás értékhez számos D-galaktóz – cAMP-CRP koncentrációpár tartozhat. A 4.7 ábrán a PI_{galE} példáján bemutatom, hogy egy bizonyos mért promóter aktivitás érték (piros pöttyök által kirajzolt sáv) számos különböző cAMP-CRP és D-galaktóz koncentrációhoz tartozhat.



4.2 ábra. A D-galaktóz és cAMP-CRP szintek meghatározásának problémája. Az általunk mért promóter aktivitás, azaz Z-tengelyen vett adat (melyet piros pöttyök alkotta sáv jelez az ábrán) számos helyen illeszkedik a promóter aktivitás modellezett felszínére, így önmagában nem elegendő a D-galaktóz és cAMP szintjének meghatározásához. A görbe palástja itt azért készült szürke skálán, mert a hatást kialakító represszor ebben a problémában irreleváns.

Több promóter egyidejű aktivitásának ismeretében számos olyan cAMP-CRP és D-galaktóz koncentrációpárt ki lehet zárni, ami ugyan az egyik promóter aktivitást kielégíti, de egy vagy több másikat nem. A rendszer promóter aktivitásainak időbeni lefutása és az azonos időpontban vett aktivitások ismeretében felállítható egy olyan modell a cAMP-CRP és a D-galaktóz szintjének alakulására, ami a legjobban tükrözi a mért promóter aktivitások időbeni lefutását, azaz a legjobban kielégíti az összes promóter aktivitás esetén vett szignál szinteket.

A probléma megoldására az előző differenciálegyenletek és a promóter aktivitások mért értékeinek felhasználásával a 4.3 folyamatábrán bemutatott módon működő algoritmust használtunk.



4.3 ábra. A simulated annealing iteratív algoritmusának folyamatábrája.

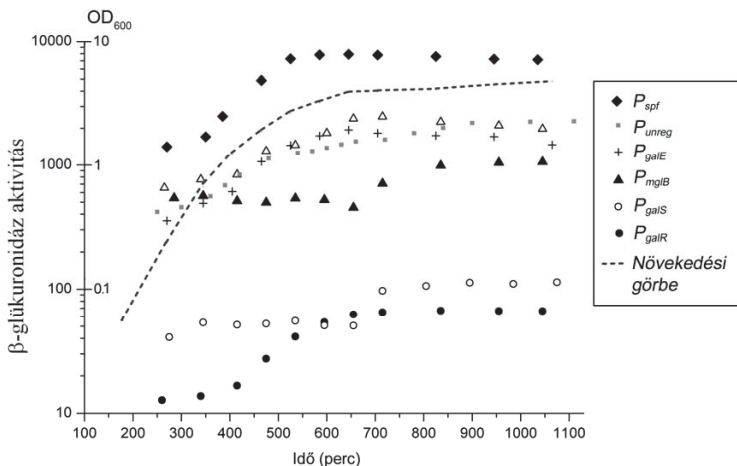
5. Eredmények

5.1 A *gal* rendszer időzítéséhez kapcsolódó eredmények

Az *E. coli* galaktóz hálózatának választát akartuk végig követni olyan esetben, mikor a sejteknek csak véges mennyiségű D-galaktóz áll rendelkezésükre véges térfogatban, ezért a λ -red rendszer segítségével létre hoztunk hét promóter-*uidA* riportert fúziós konstrukciót a *P_{galR}*, a *P_{galE}*, a *P_{galP}*, a *P_{galS}*, a *P_{mglB}* és a *P_{spf}* promóterek felhasználásával, valamint egy, a GalS, a GalR és a cAMP-CRP által nem szabályzott kontroll promóter segítségével, melyet „nem szabályozott” promóternek nevezek. A hét riportert törzset külön-külön MG1655 referenciatörzsből kiindulva hoztuk létre úgy, hogy a megfelelő promotórokra cseréltük az *uidA* gén promóterét.

A riportert törzseket LB tápoldatban növesztettük overnight, majd M63 minimál táptalajba oltottuk őket, melyben D-galaktóz volt az egyedüli szénforrás. A növekedésüket 16 órán át követtük, rendszeres időközönként mintákat véve a kultúrából, és minden esetben rögzítettük a mintavétel időpontját a kezdeti hígítástól számítva, és a szuszpenzió 600 nm-en mért optikai denzitását (OD₆₀₀). A minták 500 μ l-nyi, OD₆₀₀ = 0,2 értékre hígított mennyiségén β -glükuronidáz aktivitást mértünk. Mivel a méréshez használt sejtek száma konstans, és az enzim stabil, így az enzimaktivitás két dologtól függ: az enzim újonnan történő képződésétől és a sejtosztódás miatti hígulásától.

Az általunk tanulmányozott riportertfúziók hasonlóan működtek abból a szempontból, hogy mindegyik esetben egy kezdeti alacsony intracelluláris β -glükuronidáz szintről egy magasabb szintre jutott a sejt (5.1 ábra). A változás a promótertől függően 2-6-szoros volt. A változás időzítése és az alacsony és magas állapot közötti átmenet sebessége volt az, ami a különböző promóterektől függően szignifikánsan változott. Négy esetben (*spf*, *galE*, *galP* és *galR*) az átmenet a 400-600 perces időintervallumban történt, amikor a sejt a gyors növekedés állapotából a nagyon lassú növekedés állapotába jutott. A másik két esetben (*galS*, *mglB*) az átmenet később történik, amikor a sejtosztódás már szinte elhanyagolható (650-800 perc). A nem szabályzott promóter esetében az alacsony és magas állapot közti átmenet kevésbé éles.



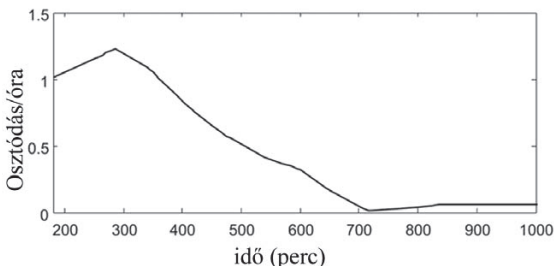
5.1 ábra: A sejtek β -glükuronidáz tartalma az idő függvényében. A promóter-*uidA* fúziót tartalmazó baktériumtörzset D-galaktóz tartalmú M63 minimál táptalajon növesztettük. A különböző időpontokban vett minták OD_{600} értékét rögzítettük, és β -glükuronidáz aktivitását megmértük. Az enzymaktivitást 0,5 ml, $OD_{600} = 0,2$ értékre hígított sejtől mértük, és az idő függvényében ábrázoltuk (három párhuzamos baktérium kultúrából mért aktivitások átlaga).

5.1.1 A növekedési görbe kiszámítása

A mért OD értékekből az alábbi képlettel meghatározhatjuk, hogyan változik a baktériumok növekedési rátája (γ) az idő (t) függvényében.

$$\gamma(t) = \frac{d}{dt} \ln(OD_t)$$

Az adataink szerint a sejtek növekedési rátájának maximuma 1,225 osztódás volt óránként $t = 285$ percnél, ami körülbelül óránként 0,011 duplázódás értékre csökkent a 715. percnél, majd ismét enyhén növekedett óránként 0,06 duplázódásig, amit a kísérlet végéig meg is tartott. A 180. perc előtt nem mértünk OD-t, de feltesszük, hogy a növekedési ráta kezdeti értéke 0, és az is marad egy körülbelül 100 perces lag fázis során (5.2 ábra.)



5.2 ábra: A növekedés görbe az idő függvényében az 5.1 ábrán bemutatott kísérletről.

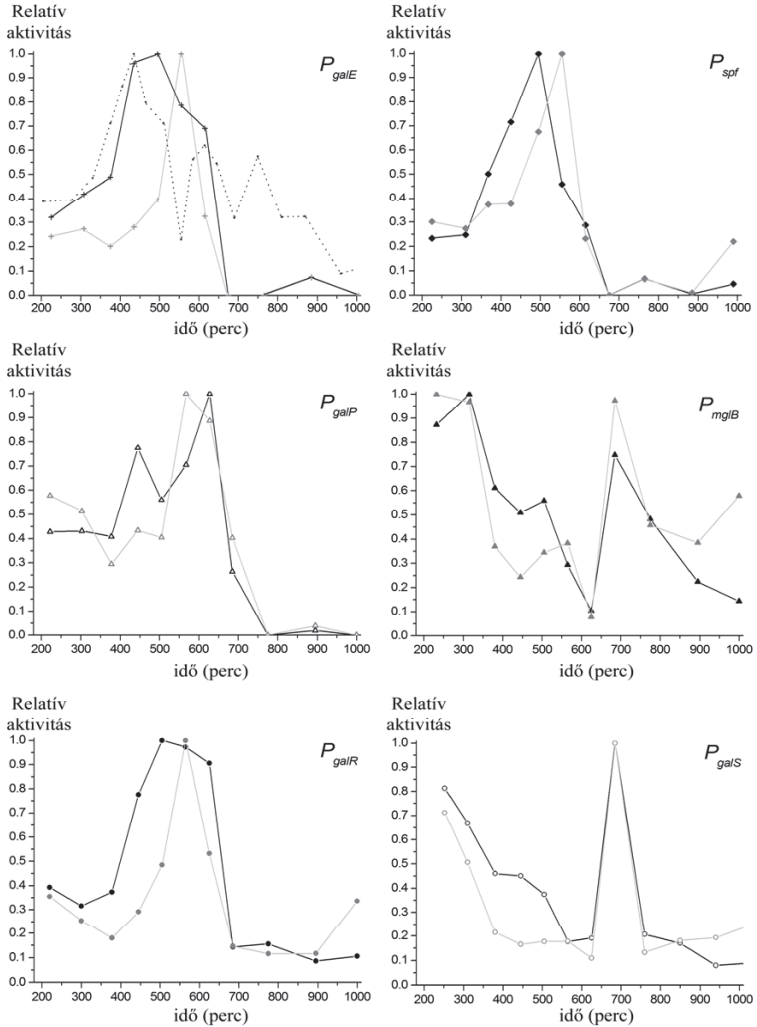
5.1.2 A promóter aktivitások kiszámítása

A β -glükuronidáz fehérje koncentrációját a sejtben (P) a fehérje termelődése és a sejt tartalmának a sejtosztódás általi hígulása határozza meg. Feltesszük, hogy a fehérje termelése arányos a promóter aktivitással (A), és a sejt kultúra optikai denzitása egy adott t időpontban (OD_t) arányos a sejtek számával, ekkor kiszámolhatjuk a promóter aktivitás időbeni alakulását az alábbi egyenlettel:

$$A = \frac{P_{t+\Delta t} - P_t \left(\frac{OD_t}{OD_{t+\Delta t}} \right)}{\Delta T}$$

A „nem szabályozott” promóter esetén mért enzimaktivitásokból számolt promóter aktivitás időbeni lefutásához – melyet 1-re normalizáltunk – viszonyítottuk a többi promóter aktivitást azért, hogy a sejtbeli források elérhetőségét (azaz a minden promóterre ható nem specifikus hatásokat) számításba tudjuk venni. A növekedési görbe és a „nem szabályozott” promóter kiszámolt aktivitása nem mutat közvetlen korrelációt, amire egy lehetséges magyarázat lehet az, hogy az ATP-képzésen kívül a galaktóz hasznosítási útvonal más sejt folyamatokat is érint. Például az elérhető UDP-galaktóz mennyisége befolyásolja a peptidoglikán szintézist és a transzkripciót. [27]

Az eredmények azt mutatják, hogy a galaktóz rendszerben a promóterek aktivitása különböző időzítési mintázatokat követ, és a „nem szabályozott” promóterhez mért (skalázott) aktivitások két különböző görbével írhatóak le. (5.3 ábra)

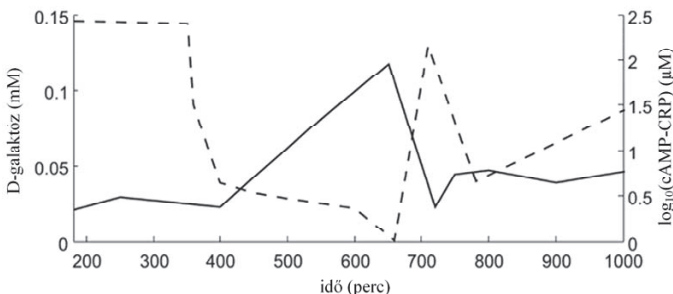


5.3 ábra. A *gal* regulon hat promóterének számolt aktivitása. A nem szabályozott promóter aktivitása, melyet az összes többi promóter skálázására használtunk, hogy a nem specifikus hatásokat számításba vehessük, a bal felső panelen szaggatott vonnal látható. A skálázott értékeket a szürke görbék mutatják. A promóterek aktivitása a saját megfigyelt maximumukat egységnyinek véve egyenként normalizálva van.

A P_{galR} , P_{galE} és a P_{spf} aktivitása gyors kezdeti növekedéssel jellemezhető, majd egy meredek esés következik, ami egy egyszeres csúcsot eredményez. A P_{galR} és a P_{galE} esetében a csúcs szélesebb a skálázotthoz képest. A skálázott P_{galP} , P_{gals} és P_{mglB} esetében az expressziós mintázat hasonló, mindegyiknek két csúcsa van. Az első a 250. perc környékén, a második a 600. percnél a P_{galP} esetében, és a 700. percnél a P_{gals} és a P_{mglB} esetében.

5.1.3 Az intracelluláris cAMP és D-galaktóz szint számítása

A mért promóter aktivitások az intracelluláris cAMP és/vagy D-galaktóz szint, és a sejt génextpressziós kapacitásának függvényei. Ebből következően a promóter aktivitások idősorai alapján vissza tudunk következtetni a cAMP és a D-galaktóz sejten belüli szintjének időbeli alakulására. A ref[27] által közölt modell leírja, hogy különböző intracelluláris cAMP-CRP és D-galaktóz szintek mellett hogyan alakul a *gal* regulon promótereinek aktivitása. Ezt a modellt használva megkerestük azt az idősort e két molekula szintjének leírására, ami a legjobban illeszkedik az általunk mért adatokhoz (5.4 ábra).



5.4. ábra. A „simulated annealing” módszerrel számított intracelluláris cAMP és D-Galaktóz szintek.

Modellünk szerint a D-galaktóz intracelluláris szintje $t = 400$ percnél kezd jelentősen növekedni, a csúcsa $t = 650$ percnél van. Ezután gyorsan csökkenni kezd, és majdnem nullára esik $t = 720$ percnél, ez ugyanaz az időpont, amikor a növekedési ráta a legalacsonyabb. Ezután ismét növekszik a D-galaktóz sejten belüli szintje, ami egybe esik az *mglBAC* rendszer 700. percen történő bekapcsolásával. Az előzetes várakozásnak megfelelően a cAMP-CRP szintje fordítottan függ össze a D-galaktóz szintjével [45], bár,

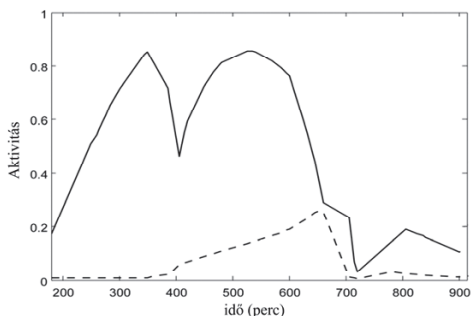
mivel a D-galaktóz az ún. nem-PTS cukrok közé tartozik, ennek a biokémiai háttere még ismeretlen.

A cAMP-CRP szintje magas értékről indul (ami nem annyira klasszikus „éhezést”, mint inkább a sejtek állandó magas igényét jelzi a cukoranyagcsere termékei iránt), és a $t = 400$ perces időpontra csökken le a szintje, majd a 670. percben ismét emelkedni kezd, és a D-galaktóz második csúcsának felfutásával egy időben, $t = 720$ percnél lecsökken.

5.1.4 Változások a $P1_{galE}$ és a $P2_{galE}$ transzkripciójában

A fentebb említett modell [27] lehetővé teszi azt is, hogy a *galETKM* transzkripcióját felbontsuk a $P1_{galE}$ és a $P2_{galE}$ promóterek aktivitásának összegére. Emlékeztetőül: az irodalmi adatok alapján a $P1_{galE}$ -ről történő transzkripció a *galETKM* operon ekvimoláris expresszióját teszi lehetővé, míg a $P2_{galE}$ -ről diszkoordinált expresszió történik, a promóterhez közel eső gének (*galE*, *galT*) javára, ami a bioszintetikus igények kielégítését szolgálja.

Kísérleteink során azt találtuk, hogy a növekedési ciklus jelentős részében főleg a $P1_{galE}$ járul hozzá a *galETKM* transzkripciójához, míg a $P2_{galE}$ aktivitása csak egy rövid időre nő meg, $t = 670$ perc időpont környékén, mikor a cAMP-CRP szint eléri a minimumát, vagyis a sejtek növekedése során ekvimoláris mennyiségben keletkeznek a *galETKM* operon enzimei, és a bioszintetikus útvonalnak megfelelő diszkoordináció csak a növekedés leállásakor történik meg (5.5 ábra).

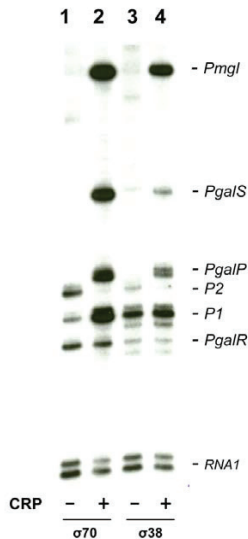


5.5. ábra A $P1_{galE}$ (folytonos vonal) és a $P2_{galE}$ (szaggatott vonal) promóterek aktivitása, úgy normalizálva, hogy a teljes aktivitás maximuma egységnyi legyen.

5.1.5 A gal regulon promótereinek transzkripciója a stacionárius fázisra jellemző (σ^{38}) RNS-polimerázzal

Az exponenciális fázisból stacionárius fázisba történő átmenet egyik velejárója az RNS-polimeráz (RNAP) σ^{70} alegységének kicserélődése σ^{38} alegységre. Ezt a folyamatot az Rsd protein segíti elő, ami lekötí a σ^{70} alegységeket. [75]

Az már korábban is ismert volt, hogy a vegetatív (σ^{70}) és a stacionárius fázisra jellemző (σ^{38}) RNS-polimeráz eltérő preferenciával írja át a *galETKM* operon két promóterét, a *P1_{galE}*-t és a *P2_{galE}*-t. [76] Fenti eredményeink azt mutatták, hogy a *P_{mglB}* promóter a növekedés későbbi fázisában aktív, a maximális expresszió 45%-át a 800. percnél mutatja. Ez felveti azt a kérdést, hogy a σ^{38} RNAP vajon a σ^{70} RNAP-hoz mérhető mértékben képes-e átírni a *P_{mglB}* promótert, ezért megvizsgáltuk a gal regulon promótereinek transzkripcióját σ^{70} és σ^{38} RNS-polimerázokkal *in vitro*, cAMP-CRP jelenlétében és hiányában.



5.6. ábra: A gal regulon promótereinek σ^{38} RNS-polimeráz általi transzkripciója. Az *in vitro* transzkripciót supercoil állapotú pRPGSM DNS templárról, HU és cAMP jelenlétében végeztük. A CRP jelenlétét vagy hiányát +/- jelzi az ábrán a minták alatt. Az 1-2 lane-ben látható kísérletekben σ^{70} RNAP holoenzimet használtunk, a 3-4 lane esetében rekonstruált σ^{38} RNAP-t.

Azt találtuk (5.6 ábra), hogy cAMP hiányában mindkét σ -faktor esetében csak a $P_{I_{galE}}$, $P_{2_{galE}}$ és a P_{galR} promóterekről történik átírás. Az eredményeink megerősítik azt a korábbi megfigyelést, hogy σ^{38} RNAP erős preferenciát mutat a $P_{I_{galE}}$ irányába, és a $P_{2_{galE}}$ -t alig írja át, ezzel szemben a σ^{70} RNAP fordított preferenciát mutat.[76] A P_{galR} , hasonlóan a $P_{2_{galE}}$ -hez, a σ^{38} RNAP esetében mutatott alacsonyabb aktivitást.

Négy promóter esetében növelte a cAMP-CRP komplex a transzkripció mértékét, σ^{38} és σ^{70} RNAP esetén is. Ezek a $P_{I_{galE}}$ (aminek cAMP-CRP hiányában is volt aktivitása mindkét σ -faktor esetén), továbbá a P_{mglB} , a P_{galP} és a P_{galS} (ezeknek cAMP-CRP nélkül nem volt aktivitásuk). Ha promóterenként összevetjük cAMP-CRP jelenlétében elért aktivitás mértékét σ^{38} és σ^{70} RNAP estében, akkor szignifikáns különbségeket tapasztalunk. Míg P_{mglB} esetében a σ^{38} RNAP általi aktivitás eléri a 70%-át a σ^{70} RNAP általi aktivitásnak, addig a P_{galS} esetén ez mindössze 20%. A P_{galS} -nek ez a gyenge σ^{38} RNAP általi transzkripciója és az erős autoregulációja magyarázatául szolgálhat a P_{galS} aktivitás gyors esésének a P_{mglB} -hez viszonyítva az 700-750 perc közötti időszakban.

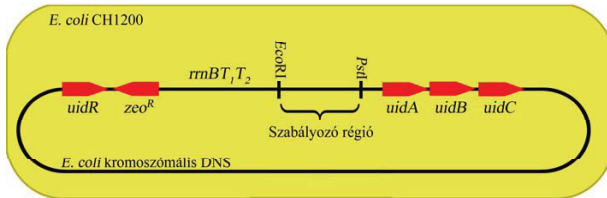
A $P_{I_{galE}}$ -t cAMP hiányában a σ^{38} RNAP írja át erősebben a σ^{70} RNAP-hoz viszonyítva, de a cAMP jelenléte ezt csak kis mértékben növeli. Ugyanakkor a σ^{70} RNAP ezt a promótert cAMP hiányában alig, jelenlétében viszont igen erősen írja át, azaz a cAMP általi aktiválódás mértéke igen nagy. Mivel a *galR* és *galS* gének csak nagyon gyengén íródnak át a stacionárius fázisra jellemző σ^{38} RNAP által, ezért úgy véljük, hogy a *gal* regulon promótereinek (főként a $P_{I_{galE}}$ -nek és a P_{mglB} -nek) a stacionárius fázisban történő repressziójáért főként az addig felhalmozódott represszor fehérlék felelősek.

Munkánk során azt találtuk, hogy a szignál integráció módja (Boole-logikai kapu) és az expresszió időzítése összefüggést mutat.

5.2 A *gal* rendszer genetikai flexibilitásával kapcsolatos eredmények

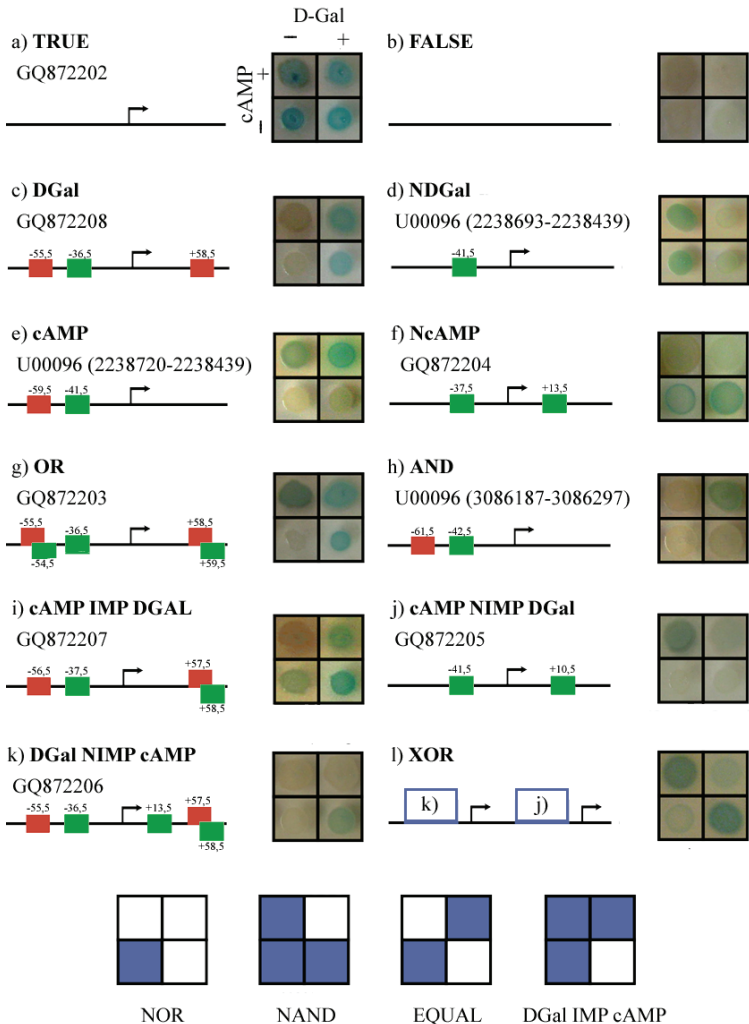
A galaktóz hálózat promóterei esetében a cAMP és galaktóz szignálok különböző módon kerülnek integrálásra. Kísérleteink fontos iránya volt a rendszer flexibilitásának vizsgálata, azaz hogy az *E. coli*-ban létező promóterek és regulációs szekvenciák mutációival minél több (törekvéseink szerint minden) logikai kaput létrehozunk. Ehhez mesterségesen előállított és a gal rendszerben létező természetes promóter konstrukciókat is létrehoztunk, s ezeket a fentebb leírt módszerrel rekombináltattuk az *uidA* gén promóter régiójának helyére (5.7 ábra). Ezeknél a kísérleteknél a kiinduló baktérium törzs a kanonikus

laboratóriumi *E. coli* törzs, az MG1655 helyett annak adenilát-cikláz (*cyaA*) deficiens mutánsa, a CH1200 volt. Erre azért volt szükség, hogy kontrollálhatóan adagolhassuk a cAMP-t a rendszerbe, s ilyen módon felismerjük a cAMP-re reagáló és nem reagáló promóter konstrukciókat.



5.7 ábra. A szignál integrációs tesztekhez használt riporter konstrukciók sematikus ábrája. A piros nyilak a nyitott leolvasási kereteket jelölik. Az ábra feltünteti az *uidABC* operon összes tagját, melyek közül az *uidB* glükuronid GPH transzportert kódol (ez a GPH, azaz galaktozid-pentóz-hexuronid transzporter család tagja), az *uidC* egy a transzportban részt vevő fehérjét kódol; valamint az operon represszorát (*uidR*) is. A szabályozó szekvenciákat az *EcoRI* és a *PstI* restrikciós helyek közé illesztettük. A *zeo^R* kazetta a konstrukció része volt, az *uidR* irányából esetlegesen érkező transzkripciót az *rrmBT₁T₂* terminátor csoport állítja meg.

A létrehozott sejtekben a szignál integrációt populációs szinten értékeltük, négy különböző LB-lemezre cseppentve a törzseket, a négy különböző szignál kombinációnak megfelelően (D-galaktóz, cAMP, mindkettő, egyik sem), az X-Gluc nevű, színreakciót adó szubsztrát jelenlétében. Ezekon a lemezen megfigyeltük a különböző konstrukciók szignálintegrációs viselkedését. A Boole-logikai kapu kritériuma az volt, hogy a magas és alacsony expressziós állapotok, azaz a telep kék és fehér volta könnyen megkülönböztethetőek legyenek, habár megjegyzendő, hogy a „magas szint” a különböző szignálkombináció esetében nem biztos, hogy ugyanazt jelenti. Munkánk során 40 különböző konstrukciót hoztunk létre olyan módon, hogy átrendeztük és módosítottuk a rendszerben előforduló transzkripciósi faktorok kötőhelyeit, ezáltal módosítva a promóterek alapaktivitását. Ezzel a módszerrel elkészítettünk a 16 lehetséges szignál integrációs logikai kapu közül 11-et. A 40 konstrukcióból 11-nek meghatároztuk a szekvenciáját (minden Boole-kapuzhoz egyet-egyét), és a kapott szekvenciát elhelyeztük a GenBank-ban. Az 5.8 ábra mutatja be a konstrukciókat, azok szignálintegrációs logikáját és feltünteti az egyes konstrukciók GenBank kódját. A FALSE kapu szekvenciája (nem tartalmaz promótert) a következő: 5'-GAGCTCGGTACCCGGGGATCGATCCTCTAGATCGAC-3', ezt nem helyeztük el a GenBank-ban.

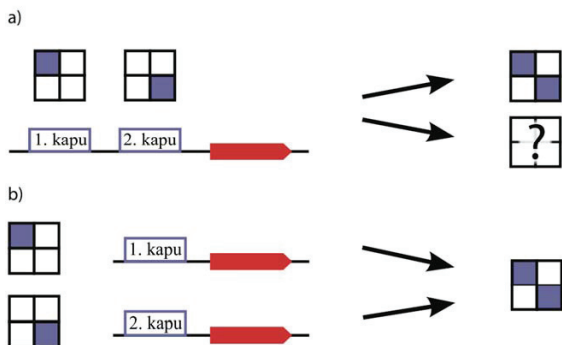


5.8 ábra. Az általunk létrehozott logikai kapuk. A piros téglalapok a GalR/GalS operátort, a zöld a CRP-kötőhelyet jelöli (nem méretarányos), a számok ezek helyett mutatják a transzkripció stratégia pontjához képest. A szabályozó régiók szekvenciáját elhelyeztük a GenBankban, az ábrán látható számok alatt érhetőek el. Alul annak a négy kapunak a sematikus ábrája látszik, amit nem sikerült létrehozunk.

A lehetséges kapuk közül hat olyan van, ami vagy csak egyféle szignálra reagál (D-Gal, ND-Gal, cAMP, NcAMP), vagy egyáltalán nem válaszol a szignálokra (TRUE, FALSE). A többi logikai kaput – amelyeknél tehát a két szignál ténylegesen integrálódik – két csoportba oszthatjuk: azokra, melyeknél a promóternek van alapaktivitása (a Boole-kapu ábráján a jobb alsó négyzet színes, ezek: OR, XOR, cAMP IMPD-Gal, NAND, D-Gal NIMP cAMP), valamint azokra, ahol nincs alapaktivitás (a jobb alsó négyzet fehér: AND, EQUAL, cAMP NIMP D-Gal, D-Gal IMP cAMP, NOR). Mivel a második csoport esetében az RNS-polimeráz (RNAP) aktív toborzására van szükség, az ide tartozó kapukat nem lehet pusztán a represszorkötőhelyek felhasználásával létrehozni. Ugyanakkor minden kaput elkészíthetünk, ha legalább egy aktivátorhelyet is használunk. Az aktivátorok sokoldalúbbak a represszoroknál, mivel ezek tipikusan kettős funkciójú szabályozó fehérjék: a promóter központi régiójába kötött aktivátorok gátolhatják az RNAP bekötődését, így csökkenthetik a transzkripció szintjét. Az RNAP aktivátorok és represszorok általi szabályozott toborzása a transzkripció szabályozásának leggyakoribb módja [77], amihez jól illeszkedik az a megfigyelésünk, hogy aktivációs és repressziós mechanizmusok kombinációjával szabályozási logikák igen széles skáláját tudtuk megvalósítani.

5.2.1 Logikai kapuk kombinálása

Genetikai logikai kapuk kombinálásával újabb kapukat állíthatunk elő. Ennek két módja lehet, melyeket párhuzamos és soros kapcsolásnak nevezünk (5.9 ábra).



5.9 ábra. A kapuk kombinálásának két stratégiája. A soros kapcsolás esetén (a) a szekvenciákat egymás után fuzionáltatjuk a riportergén upstream régiójába, a kettő együttes logikája lehet a két külön logika összege, de lehet valami teljesen eltérő logika is. A párhuzamos kapcsolásnál (b) a szabályozó régiók külön-külön vannak a riportergén elé fuzionáltatva, és eredményként a kapuk összegét várjuk.

A párhuzamos felépítés során a két logikai kaput külön-külön fuzionáltatjuk a riportergén elé. Ez a felállítás a génduplikációt követő, a szabályozó elemek mutációja miatt bekövetkező expressziós divergencia hatását másolja.[78] Mivel a két kópia egymástól függetlenül expresszálódik, ilyen esetben az összességében tapasztalható szabályozási logika a két kópia logikájának összege lesz, pl a „cAMP” és a „D-Gal” kapuk párhuzamos módon kapcsolt összege egy „OR” kapu lesz.

A soros kapcsolás esetében a két logikai kaput (szabályozó szekvenciát) ugyanazon riportter génnel fuzionáltatjuk, ami analóg azzal, mikor egy gén több promotérről is átíródhat, mely promotéreket ugyanazok a transzkripciósi faktorok szabályozzák (de eltérő módon). A munkánkban azt teszteltük, hogy a soros kapcsolási módokból hogyan jöhetnek létre új logikai kapuk. Megpróbáltuk létrehozni az „XOR” („D-Gal NIMP cAMP” + „cAMP NIMP D-Gal”), a „NAND” („D-Gal NIMP cAMP” + ND-Gal), valamint a „D-gal IMP cAMP” („AND + ND-Gal) kapukat. Hat tervezett konstrukcióból azonban csak egy adta a várt XOR funkciót, a többi esetben nem Boole-típusú integrációt figyeltük meg, vagy az egyik, a konstrukcióban részt vevő kapu dominált a többi szabályozóelem felett.

6. Diskusszió

Munkánk során az *Escherichia coli* galaktóz hasznosítási rendszerén végeztünk vizsgálatokat. A feast-famine ciklus modelljeként a baktériumokat nagy, de mégis limitált mennyiségű galaktózon növesztettük, és végig követtük a rendszer promótereinek időbeli viselkedését. Vizsgáltuk továbbá a genetikai elemek kombinációjában rejlő lehetséges logikai kimeneteleket.

A feast-famine ciklust modellező kísérletben azt találtuk, hogy a galaktóz metabolizmus egyrészt globális hatással van a géneexpresszióra (például mert befolyásolja a nukleotid-trifoszfát szintet), másrészt vannak specifikus hatások, amik a regulon egyes promótereinek transzkripcióját érintik (ezeket a D-galaktóz és a cAMP sejten belüli szintje befolyásolja a GalS, GalR és CRP fehérjék közvetítésével).

Munkánk során kiszámoltuk a promóter aktivitások időbeli változásain kívül az azokat kísérő (vagy előidéző) D-galaktóz és cAMP-CRP szinteket. A sejt galaktóz rendszere az intracelluláris galaktóz és cAMP szintet valós időben figyeli, és ennek megfelelően állítja be a promóter aktivitásokat. Mivel a *gal* rendszer fehérjéi stabilnak tekinthetők (hasonlóan a használt riporterfehérjéhez), egyedül a fehérjék termelésével lehet a rendszer állapotán változtatni, vagyis a rendszerben a döntések ezt fogják érinteni. Az aktuális fehérje mennyiség a fehérje termelés történetén (azaz a korábbi promóter aktivitásokon), és a híguláson (a sejosztódás mértékén) múlik. Munkánk során azt találtuk, hogy egy adott promóterről történő expresszió az aktuális sejten belüli szignálszintektől (cAMP és D-galaktóz) múlik, amit pedig a sejten kívüli galaktóz szint és a galaktóz transzportot és felhasználást aktuálisan befolyásoló tényezők (fehérjék és a Spot42 sRNS) határoznak meg.

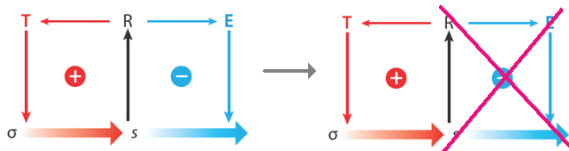
Mindezek alapján a rendszer döntései inkább a sejt mindenkori állapotán múlnak helyett, hogy a sejt az extracelluláris tér mindenkori állapotát figyelné.

6.1 A galaktóz transzport és a felhasználás koordinálása

A *galETKM* operon által kódolt enzimek a galaktóz amfibolikus felhasználásáért felelősek. Mindegyikük szükséges a galaktóz metabolizmushoz, ám csak a GalE szükséges ahhoz, hogy extracelluláris galaktóz híján a bioszintetikus glikozilációs reakciókhoz szubsztrátot szolgáltatson. A PI_{galE} promóterrel történő átírás az operon enzimeinek

ekvimoláris termelését teszi lehetővé, ami a katabolikus szükségleteknek felel meg, a $P2_{galE}$ -től diszkoordinált átírás történik, mely a promóterközeli fehérjék génjeit ($galE$, $galT$) részesíti előnyben a bioszintetikus szükségletekhez. Az az eredményünk, hogy a $P2_{galE}$ aktivitása elhanyagolható akkor, ha galaktóz batch kultúrán növesztjük a sejteket, egybeeséssel ezzel a modellel.

Ugyanakkor a fehérjék termelődése poszt-transzkripciósan is diszkoordinálódhat, a Spot42 kis RNS ugyanis a GalK termelődését gátolja cAMP-CRP hiányában[1]. Ennek promótere, a P_{spf} egy időben expresszálódik a D-galaktóz intracelluláris szintjének növekedésével és a cAMP szint esésével. Azt gondoljuk, hogy a GalK termelődésének ez a gátlása részt vesz a transzport és a metabolizmus közti egyensúly megteremtésében, és ez a szabályozás egy olyan környezetben előnyös, ahol az extracelluláris galaktóz szint csökken. Az extracelluláris galaktóz szint csökkenésével a transzport korlátozottan működik, ezt a sejt a transzporter fehérjék koncentrációjának növelésével kompenzálja. Ahhoz, hogy nagyobb mértékben expresszálhasson transzporttereket, szükség van egy magasabb intracelluláris galaktóz szintre, de ez egyben a felhasználásért felelős gének megnövekedett termelésével is együtt járna, ezt szemlélteti a 6.1 ábra bal oldala (ami egyidejűleg a bevezetőben bemutatott 2.2 ábra részlete), ami a rendszert mint két összekapcsolt visszacsatolási kört mutatja be. Ha a sejt nem gátolná a galaktóz fogyásával (csökkent influx) egy időben a felhasználást, akkor még azelőtt fogyna el az intracelluláris galaktóz, hogy a sejt elegendő „kompenzáló” transzportert tudott volna termelni, és így a represszorok még ez előtt kikapcsolnák a rendszer génjeit. Ezt a hatást lehet ellensúlyozni azzal, ha a sejt megváltoztatja a két visszacsatolási kör arányát, azaz szétkapcsolja azokat.[7]



6.1 ábra. A két visszacsatolási kör szétkapcsolása (a felhasználási kör csökkent működése) a cukor felhalmozódásához vezet a sejtben belül.

A GalK termelés korlátozása hatékony módja annak, hogy a sejt blokkolja a metabolikus visszacsatolási kört a transzport veszélyeztetése nélkül. Ugyanakkor a Spot42 nagy arányban képződik akkor, ha a cAMP szint alacsony, függetlenül attól, hogy az extracelluláris D-galaktóz szint csökken-e vagy sem. A transzportkapacitás növelése nem

szükséges galaktózban gazdag környezetben, ezért azt gondoljuk, hogy a transzport és metabolikus kör poszttranszkripció szétkapcsolása a fluktuáló galaktózkoncentrációhoz való adaptáció eredménye. A Spot42 termelése lehetővé teszi, hogy a sejt hatékonyabban hasznosítsa a galaktózt, ha annak extracelluláris szintje alacsony, és mivel a Spot42 nagy mennyiségben van jelen, hatása több generáció során is fennmarad, még ha a szintézise le is állt.[55]

Ezek alapján azt gondoljuk, hogy az *E. coli* galaktóz rendszere egy olyan környezetben való működéshez adaptálódott, melyben a galaktóz szintje időben változik, azonban a válaszadási mechanizmus jobban illik olyan környezetekre, melyekben a jövőbeni galaktóz szint nem jósolható, mint a szabályos feast-famine ciklusokra, mely utóbbiban a fluktuációk szabályosak, azaz adaptálódni lehet hozzájuk. Ezt mutatja a $t = 700$. percnél bekövetkező majdnem 0 növekedési ráta (ez egybe esik a cAMP csúccsal), amit a sejt kivédhetne azzal, ha az MglBAC transzportrendszert akkor kapcsolná be, mikor az intracelluláris D-galaktóz szint csökkenni kezd – ezt várnánk akkor, ha a sejt a hagyományos feast-famine rendszerhez optimalizálta volna a gal rendszert.

6.2 A galaktóz rendszer promótereinek transzkripció mintázatát meghatározó tényezők

A *gal* regulon promótereinek transzkripciója bármely adott időpontban a (ld 5.1 ábra) a következőktől függ:

- A transzkripcióhoz szükséges források elérhetősége.
- Az intracelluláris cAMP és D-galaktóz szintek kombinációja.
- A szignálok érzékelő regulátor fehérjék koncentrációja.
- Az adott promóter és szabályozó régió struktúrája és funkciója.

Ezt megelőzően leírták, hogy milyen módon függ a *gal* regulon promótereinek transzkripciója a cAMP és D-galaktóz koncentrációktól, és azt találtuk, hogy a két szignál hasonlít a Boole-logikai függvényekre. [11, 27] Jelen munkánkban azt találtuk, hogy a skálázott promóter aktivitás görbék alakja (ld. 5.3 ábra) és a promóter szignál integrációs logikája közt korreláció áll fenn. Egy aktivitási csúcsot figyeltünk meg a DGal és a NOT cAMP logikájú promótereknél, és két csúcsot az AND-szerű kapuknál. Az utóbbi esetben a csúcsok helyzete attól függ, hogy milyen erősen köti a GalR/GalS és a cAMP az

adott promóter szabályozó régióit. Például a P_{galP} esetében, ami a legerősebb GalR kötőhelyet tartalmazza, és egyben a legérzékenyebb cAMP-CRP általi aktivációra, a második aktivációs csúcs korábban történik, mint a P_{gals} esetében, ami csak jóval magasabb cAMP-CRP szinttel aktiválható, és aminek GalR/GalS általi represszióját kevesebb D-galaktóz is fel tudja oldani.

6.3 A szabályozási hálózat flexibilitása

A lehetséges szignálintegrációkat vizsgáló munkánk során a génszabályozás legegyszerűbb kölcsönhatásait használtuk fel: az RNAP-promóter interakciót, a transzkripció faktor (TF) -kötőhely interakciót és a TF-szignálmolekula interakciót. Eredményeink azt sugallják, hogy a baktériumok szabályozási hálózatai nagyon rugalmasak lehetnek a szignálintegráció tekintetében. Létrehoztunk számos különböző szabályozó szekvenciát, melyek közül szekvenáltunk és annotáltunk 12 különböző Boole-logikai választ adó szekvenciát. Elvileg a maradék négy kapu is létrehozható a már létező moduljainkból. Például a „NOR” kapu létrehozható lehetne egy cAMP-CRP által szabályozott NDgal konstrukcióból, és utána az összes többi konstrukciót megalkothatnánk a megfelelő párhuzamos kombinációk segítségével (azaz: EQUAL = NOR+AND; DGal IMP cAMP = NOR + cAMP; NAND = NDgal + NcAMP). Habár kísérleteink során 12 különböző logikai kaput sikerült létrehoznunk, számos esetben nem értjük pontosan a szabályozó régió struktúrája és funkciója közti összefüggést. A legérdekesebb az NDgal konstrukció (5.8 ábra), ahol nem tudtunk GalR-kötőhelyet azonosítani a szekvenálás során. Mivel a kapu értékelésekor a riporter fehérje aktivitását figyeljük, nem pedig a riportergén transzkripcióját, ezért nem zárhatjuk ki a szignálmolekulák által végrehajtott poszttranszkripció módosulást (pl riboswitch segítségével) sem.

Úgy véljük, hogy a transzkripció aktiváció, azaz az RNAP toborzása a promóterhez a szabályozási logika rugalmasságának fontos meghatározó tényezője. Alig néhány olyan logikai kapu van, ami két represszorral megvalósítható, ugyanakkor elvileg minden logikai kaput meg lehet valósítani aktivátorok segítségével. Az aktivátor funkciója könnyen variálható az RNAP toborzás és az RNAP gátlás között azáltal, hogy áthelyezzük a kötőhelyét. Bár általában az aktivátorok gyengébben kötik a DNS-t, mint a represszorok, a szterikus gátláshoz szükséges erősebb kötés elérhető mutációkkal vagy több kötőhely használatával.

A szabályozó régiókon kívül maguk a szabályozó TF-okat termelő gének is hozzájárulnak a hálózatok genetikai flexibilitásához. A szabályozási logika erősen megváltozhat, ha például a TF génje elszenved egy missense mutációt. Például a LacI represszor analízise megmutatta, hogy a pontmutációk megváltoztathatják azt, hogy a fehérje milyen erősen tudja kötni a DNS-t, mennyire képes DNS-hurkot képezni, mennyire köti a saját inducer molekuláját, sőt, még a represszor indukciós profilja is megváltozhat, ami miatt az általa szabályozott rendszer logikája is az ellenkezőjére fordulhat. (Például a normális működéssel ellentétben a TF az inducer jelenlétében kötődik a DNS-hez.) [79, 80]) A galaktóz rendszerben egy fordított indukciós profiljú GalR az NDGal kaput DGal kapuvá változtatná, vagy a NOR kaput DGal NIMP cAMP kapuvá.

6.4 A logikai kapuk kombinálása

Mind a párhuzamos, mind a soros módon kapcsolt logikai kapukra ismerünk példákat a természetben.[4, 81] A párhuzamos kapcsolás létrejöhet génduplikációval, amit a szabályozó szekvenciák divergálódása követ. A génduplikációval történő evolúció megoldhat adaptációs problémákat a szabályozó régióban. Például ha bizonyos logikai kapuk nem valósíthatók meg pontmutációkkal vagy a szabályozóelemek átrendeződésével, létrejöhetnek két egyszerűbb kapu párhuzamos kapcsolatával. Ilyenkor a részt vevő szabályozó elemek kiegészítik egymást funkcióját, így elősegítve azt, hogy mindkét kópia fennmaradjon a genomban. Az élesztőkön és magasabb rendű eukariótákon végzett genom analízis valóban azt mutatja, hogy a duplikált gének sokkal valószínűbben mutatnak divergált expressziós mintákat, mint az egykópiások – mind genomon belül, mind genomok között.[82] Viszont baktériumokban a magas rekombinációs ráta miatt a szabályozórégiók duplikációval történő evolúciója korlátozott. Így olyan sejtekben, ahol a homológ rekombináció aktívan zajlik, a soros kapcsolás sokkal jobb megoldás. És valóban számos *E. coli* gén íródik át több promóterről, de azonos TF-ok szabályozása alatt, mint amilyen a $P1_{galE} / P2_{galE}$ esetében is láthatunk. Az eredményeink azt mutatják, hogy a logikai kapuk sorba kapcsolása valóban kialakíthat új szabályozási logikát. Ugyanakkor az így létrejött új kapuk esetében az új funkció a párhuzamos kapuknál kevésbé prediktálható. Számos tényező befolyásolhatja az új szabályozó régió működését, például:

- Egy új promóter létrejött. Ha egy erős aktivátort arra használunk, hogy represszáljon egy promótert, előfordulhat, hogy a represszált promóterből downstream elhelyezkedő szekvenciák az aktivátor hatására promóterként fog viselkedni. Az is megtörténhet, hogy két szabályozó szekvencia fúziójából egy szignifikáns promóteraktivitással rendelkező DNS szekvenciát kapunk. Az ilyen új promóterek megváltoztathatják a szabályozási logikát akár a gén átírásával, akár promóter interferencia révén. [83]
- Transzkripciós roadblock („útakadály”). Ebben az esetben egy erősen kötő TF gátolhatja egy tőle upstream elhelyezkedő promóterrel az átírást. [84] Ez megmagyarázza, hogy hogyan lehet a downstream elhelyezkedő logikai kapu domináns a két kapuból álló szabályozó régió egészére nézve.
- Távolság ható kommunikáció. Számos TF képes kölcsönhatásba lépni az RNAP-zal vagy más TF-okkal, melyek a DNS távolabbi pontjához kötődnek. [62, 85] Így egy TF, aminek a kapcsolt kapuk közül az upstream levőt kellene szabályoznia, képes lehet befolyásolni a downstream levőt.

Jelen munkánkban mutattunk egy példát arra, mikor két logikai kapu soros kapcsolása a két kapu összegét hozta eredményül. (A DGal NIMP cAMP és cAMP NIMP DGal kapuk XOR-t hoztak létre.)

Bár sok esetben nem könnyű előre megmondani, hogy egy mesterségesen tervezett kapu hogyan fog viselkedni, ezért szabályozási logikák tudatos tervezése nehézségekbe ütközik, de a rendszer ilyen nagyfokú flexibilitása azt sugallja, hogy populáció szinten, véletlenszerű mutációkat feltételezve jó esély van arra, hogy egy új környezetbe kerülve, egy másfajta szabályozást igénylő rendszer jó eséllyel kialakulhat a már előtte meglévőből.

7. Összefoglalás

A szabadon élő baktériumsejtek egyik laboratóriumi modelljéről, az *Escherichia coli*ról mára már nagy mennyiségű adat áll rendelkezésünkre. Az egyes gének és termékeik elemzésén kívül egyre több adatunk van ezen géntermékek által alkotott hálózatokról, ezek kapcsolatáról. A hálózatok elemzése során több olyan motívumot írtak le, ami a különböző hálózatokban hasonló funkciót lát el, ezek közül a cukorhasznosító rendszerekben általános az egy pozitív és egy negatív visszacsatolás kör összekapcsolódásával létrejött motívum. Mindezek alapján világossá vált, hogy a baktériumsejt optimális működéséhez, alhálózatainak összehangolásához viszonylag korlátozott számú lehetséges stratégia létezik.

Az *E. coli* galaktóz hasznosítási hálózata egy részleteiben jól ismert, régóta kutatott hálózat. Szabályozása két bemenő jel alapján történik, melyek a D-galaktóz és a cAMP. A D-galaktóz hatását két egymással homológ, de működésükben jelentős eltérést mutató represszor, a galaktóz represszor (GalR), és a galaktóz izorepresszor (GalS) közvetíti a hálózat felé. A cAMP-t annak sejtbeli receptora, a CRP (cAMP Receptor Protein) érzékeli, és ez a cAMP-CRP komplex a DNS-t a megfelelő kötőhelyen kötve többnyire aktivátor funkciót lát el. A rendszernek tagja még két transzporter operon, az egycisztronos *galP* operon (kis affinitású, ATP-t nem igénylő galaktóz/H⁺-szimporter), és az *mgIBAC* operon (nagy affinitású, ATP-igényes transzporter), egy szabályozó kis RNS (Spot42), továbbá két, a galaktóz felhasználásában részt vevő operon, a *galETKM* és a *galU* operon, mely utóbbi nem áll galaktóz szabályozás alatt.

A *galETKM* operonnak két, egymással részben átfedő promotere van, a *P1_{galE}* és a *P2_{galE}*, melyeket a cAMP-CRP és a GalR eltérően szabályoz: a katabolikus igényeknek megfelelő *P1_{galE}*-t a cAMP-CRP aktiválja, a GalR gátolja, az anabolikus igényeknek megfelelő *P2_{galE}*-t a cAMP-CRP gátolja, de kis mennyiségű GalR aktiválja.

A két bemenő szignállal rendelkező rendszerek működését a Boole-függvényekkel lehet közelíteni, melyek a rendszernek a két szignál jelenlétének vagy hiányának kombinációjára adott igen-nem választát vizsgálják. Összesen 16-féle ilyen logikai kapu létezik, melyek közül a *gal* rendszer promoterei között csak néhány jelenik meg.

Munkánk során a *gal* rendszer promotereinek időzítését és a rendszer flexibilitását vizsgáltuk. Az időzítés vizsgálatához *E. coli* K-12 MG1655 baktériumtörzset, illetve annak *uidA* génjét, továbbá a rendszer hat legfontosabb promóterét és egy nem szabályozott

promótert felhasználva hoztunk létre promóter-*uidA* fúziós riportert törzseket. Megmértük a riportert enzim aktivitását a baktérium növekedésének számos pontján. Megállapítottuk a promóterek aktivitás változásának időbeni lefutását, valamint a simulated annealing módszer és egy recens matematikai modell segítségével következtettünk a D-galaktóz és a cAMP-CRP szint időbeli alakulására a sejtben. Vizsgáltuk a rendszer promótereinek a σ^{38} RNS polimeráz általi átíródását.

Eredményeink azt sugallják, hogy az *E. coli* sejtek nem a környezet állapotát monitorozzák, hanem a saját mindenkor belső állapotukat.

Megfigyeltük, a sejt a csökkenő galaktóz szintre a visszacsatolási körök szétkapcsolásával (a felhasználás és a transzport arányának változtatásával), végső soron galaktóz felhalmozással reagál.

Azt is megfigyeltük, hogy az egyes promóterek szignálintegrációs logikája és időzítése összefüggést mutat.

Munkánk során mesterségesen létrehozott és természetes szabályozó régiók és az adenilát-cikláz hiányos (*ΔcyaA*) *E. coli* CH1200 törzs felhasználásával további promóter-*uidA* fúziós konstrukciókat hoztunk létre, és vizsgáltuk ezek szignálintegrációs viselkedését. Négy kivételével előállítottunk mind a 16 különböző logikai kaput.

Az eredményül kapott kapukat elemezve megállapítottuk, hogy a rendszer igen flexibilis, mutációkkal könnyen alakítható egy szabályozási logika egy másikká, ami gyors alkalmazkodó képességet jelent. Úgy gondoljuk, hogy a transzkripció aktiváció igen fontos mechanizmus a flexibilis szabályozáshoz.

8 Summary

The function of living cells is controlled by complex regulatory networks that are built of a wide diversity of interacting molecular components. The sheer size and intricacy of molecular networks of even the simplest organisms are obstacles toward understanding network functionality. In this work we address two important network features, regulation of timing of gene expression and the flexibility of gene regulation in integrating environmental signals. We use the galactose regulatory system of *Escherichia coli* as a model system because the elements of this network are well characterized and a simplified mathematical model describing the behavior of the network has been previously published. The *gal* system of *E. coli* contains genes involved in the transport (*galP*, *mglBAC*) and amphibolic utilization (*galETKM*) of the sugar D-galactose. Genes of the *gal* regulon belong to different operons. This setup allows differential regulation of functions when needed. Regulation of the *gal* system is governed by two similar regulators, GalR and GalS, which are regulated in different ways. Besides sensing the intracellular sugar level, galactose utilization is also regulated by the cAMP-CRP complex. cAMP is a signal of carbon shortage and sensed by CRP (cAMP Receptor Protein).

Proper timing of gene expression is crucial in regulation of critical biological phenomena, i.e., cellular adaptation, differentiation, and development. Biological systems evolved two fundamentally different mechanisms for the regulation of timing. One possibility is that after an initial decision a regulatory network proceeds autonomously executing a pre-defined program. The other possibility is that the network monitors specific signals in real-time and regulates timing accordingly. In the natural environment bacterial cells have to adjust their metabolism to alterations in the availability of food sources. The goal of our work was to understand the principles governing the timing of gene expression in sugar networks when cells are grown in a batch culture containing a single sugar as a carbon source. We simulated the famine-feast cycle of bacterial growth by diluting stationary phase *E. coli* cells in fresh medium containing galactose as the sole carbon source. We monitored how promoter activities in the galactose utilization network change as the cell population grows and galactose is exhausted from the medium, using specific promoter-*uidABC* (*gus*) reporter fusions. Besides responding to system-specific regulators, gene expression is also influenced by global effects, which depend on the physiological

state of the cell. Because cell physiology changes in our experiments, we use a promoter which is not regulated by GalR, GalS or cAMP-CRP as a control. Based on the experimental results, we built a mathematical model to calculate intracellular D-galactose and cAMP-CRP levels. We found that the cell responds to a decreasing external galactose level by increasing the internal galactose level, which is achieved by limiting galactose metabolism and increasing the expression of transporters. We showed that the cell alters gene expression based primarily on the current state of the cell, and not on monitoring the level of extracellular galactose in real-time. Some decisions have longer term effects, therefore the current state does subtly encode the history of food availability. In summary, our measurements of timing of gene expression in the galactose system suggest that the system has evolved to respond to environments where future galactose levels are unpredictable, rather than regular feast and famine cycles.

The galactose network, similar to many other sugar networks, is a natural two-input system, regulated by a specific sugar (D-galactose) and the global regulatory signal cAMP. Therefore it is an optimal system to study signal integration at promoters. Gene regulation is based on simple building blocks such as promoters, transcription factors (TFs) and their binding sites on DNA. But how diverse are the functions that can be obtained by different arrangements of promoters and TF binding sites? We constructed synthetic regulatory regions using promoter elements and binding sites of two non-interacting TFs (GalR and CRP), each sensing a single environmental input signal. We showed that simply by combining these three kinds of elements, we can obtain 11 out of the 16 Boolean logic gates that integrate two environmental signals *in vivo*. Further, we demonstrated how combination of logic gates can result in new logic functions. Our results suggest that simple elements of transcription regulation form a highly flexible toolbox that can generate diverse functions under natural selection.

9. Köszönetnyilvánítás

Köszönöm témavezetőmnek, Dr. Semsey Szabolcsnak a belém fektetett munkáját és a hozzám való türelmét, valamint azt, hogy a laborjába fogadott. Köszönöm Dr. Orosz László professzor úrnak, hogy az ajtaja mindig nyitva állt, valamint Dr. Vellai Tibor tanszék vezető úrnak, hogy a tanszéken dolgozhattam az ösztöndíj lejárta után. Köszönöm Alexander Hunzikernek, hogy elmagyarázott bármit, amit a matematikai alapokból nem értettem. Köszönöm Sandeep Krishnának, Erdős Jánosnak és Tuboly Csabának a nekem nyújtott segítségüket. Hálával tartozom továbbá Takácsné Botond Juditnak az elsőrangú asszisztenciáért, valamint Dr. Holczinger Andrásnak a beszélgetéseikért.

Köszönöm feleségemnek, Koller Borinak, hogy elviselte, amíg ezt a dolgozatot megírtam, és szüleimnek, akik lehetővé tették, hogy eljussak idáig.

10. Felhasznált irodalom

1. Kolter, R., D.A. Siegele, and A. Tormo, *The stationary phase of the bacterial life cycle*. Annu Rev Microbiol, 1993. **47**: p. 855-74.
2. Monod, J., *The Growth of Bacterial Cultures*. Annu Rev Microbiol, 1949. **3**: p. 371-394.
3. Neidhardt, F.C., Umbarger, E.H., *Chemical Composition of Escherichia coli*, in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, N. F.C., Editor. 1996, American Society of Microbiology (ASM) Press: Washington, DC USA.
4. Keseler, I.M., et al., *EcoCyc: a comprehensive view of Escherichia coli biology*. Nucleic Acids Res, 2009. **37**(Database issue): p. D464-70.
5. Ma, H.W., J. Buer, and A.P. Zeng, *Hierarchical structure and modules in the Escherichia coli transcriptional regulatory network revealed by a new top-down approach*. BMC Bioinformatics, 2004. **5**: p. 199.
6. Sneppen, K., S. Krishna, and S. Semsey, *Simplified models of biological networks*. Annu Rev Biophys, 2010. **39**: p. 43-59.
7. Krishna, S., S. Semsey, and K. Sneppen, *Combinatorics of feedback in cellular uptake and metabolism of small molecules*. Proc Natl Acad Sci U S A, 2007. **104**(52): p. 20815-9.
8. Monod, J., J.P. Changeux, and F. Jacob, *Allosteric proteins and cellular control systems*. J Mol Biol, 1963. **6**: p. 306-29.
9. Weickert, M.J. and S. Adhya, *The galactose regulon of Escherichia coli*. Mol Microbiol, 1993. **10**(2): p. 245-51.
10. Adhya, S., *Suboperonic regulatory signals*. Sci. STKE, 2003. **2003**(185): p. pe22.
11. Semsey, S., et al., *Signal integration in the galactose network of Escherichia coli*. Mol Microbiol, 2007. **65**(2): p. 465-76.
12. Weickert, M.J. and S. Adhya, *Isorepressor of the gal regulon in Escherichia coli*. J Mol Biol, 1992. **226**(1): p. 69-83.
13. Weickert, M.J. and S. Adhya, *Control of transcription of gal repressor and isorepressor genes in Escherichia coli*. J Bacteriol, 1993. **175**(1): p. 251-8.
14. Geanakopoulos, M., et al., *GalR mutants defective in repressosome formation*. Genes Dev., 1999. **13**(10): p. 1251-1262.
15. Semsey, S., et al., *Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding*. EMBO J., 2002. **21**(16): p. 4349-4356.
16. Choy, H.E., et al., *Repression and activation of promoter-bound RNA polymerase activity by Gal repressor*. J. Mol. Biol., 1997. **272**(3): p. 293-300.
17. Choy, H.E., et al., *Repression and activation of transcription by Gal and Lac repressors: involvement of alpha subunit of RNA polymerase*. EMBO J., 1995. **14**(18): p. 4523-4529.
18. Roy, S., et al., *GalR represses galPI by inhibiting the rate-determining open complex formation through RNA polymerase contact: a GalR negative control mutant*. J. Mol. Biol., 2004. **344**(3): p. 609-618.
19. Choy, H.E. and S. Adhya, *Control of gal transcription through DNA looping: inhibition of the initial transcribing complex*. Proc. Natl. Acad. Sci. U. S. A., 1992. **89**(23): p. 11264-11268.

20. Irani, M.H., L. Orosz, and S. Adhya, *A control element within a structural gene: the gal operon of Escherichia coli*. Cell, 1983. **32**(3): p. 783-8.
21. Semsey, S., et al., *DNA trajectory in the Gal repressosome*. Genes Dev, 2004. **18**(15): p. 1898-907.
22. Virnik, K., et al., *"Antiparallel" DNA loop in gal repressosome visualized by atomic force microscopy*. J Mol Biol, 2003. **334**(1): p. 53-63.
23. Ryu, S., et al., *GalR-mediated repression and activation of hybrid lacUV5 promoter: differential contacts with RNA polymerase*. Gene, 1998. **223**(1-2): p. 235-45.
24. Kolb, A., et al., *Transcriptional regulation by cAMP and its receptor protein*. Annu Rev Biochem, 1993. **62**: p. 749-95.
25. Møller, T., et al., *Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon*. Genes Dev, 2002. **16**(13): p. 1696-706.
26. Semsey, S., et al., *Dominant negative autoregulation limits steady-state repression levels in gene networks*. J Bacteriol, 2009. **191**(14): p. 4487-91.
27. Krishna, S., et al., *Relation of intracellular signal levels and promoter activities in the gal regulon of Escherichia coli*. J Mol Biol, 2009. **391**(4): p. 671-8.
28. Lawson, C.L., et al., *Catabolite activator protein: DNA binding and transcription activation*. Curr Opin Struct Biol, 2004. **14**(1): p. 10-20.
29. McKay, D.B. and T.A. Steitz, *Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA*. Nature, 1981. **290**(5809): p. 744-9.
30. Ebright, R.H., Y.W. Ebright, and A. Gunasekera, *Consensus DNA site for the Escherichia coli catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the E. coli lac DNA site*. Nucleic Acids Res, 1989. **17**(24): p. 10295-305.
31. Izu, H., et al., *Differential control by IHF and cAMP of two oppositely oriented genes, hpt and gcd, in Escherichia coli: significance of their partially overlapping regulatory elements*. Mol Genet Genomics, 2002. **266**(5): p. 865-72.
32. Choy, H.E. and S. Adhya, *RNA polymerase idling and clearance in gal promoters: use of supercoiled minicircle DNA template made in vivo*. Proc. Natl. Acad. Sci. U. S. A., 1993. **90**(2): p. 472-476.
33. Paul, L., et al., *Integration of regulatory signals through involvement of multiple global regulators: control of the Escherichia coli gltBDF operon by Lrp, IHF, Crp, and ArgR*. BMC Microbiol, 2007. **7**: p. 2.
34. Gaston, K., A. Kolb, and S. Busby, *Binding of the Escherichia coli cyclic AMP receptor protein to DNA fragments containing consensus nucleotide sequences*. Biochem J, 1989. **261**(2): p. 649-53.
35. Burr, T., et al., *DNA sequence elements located immediately upstream of the -10 hexamer in Escherichia coli promoters: a systematic study*. Nucleic Acids Res, 2000. **28**(9): p. 1864-70.
36. Keilty, S. and M. Rosenberg, *Constitutive function of a positively regulated promoter reveals new sequences essential for activity*. J Biol Chem, 1987. **262**(13): p. 6389-95.
37. Mitchell, J.E., et al., *Identification and analysis of 'extended -10' promoters in Escherichia coli*. Nucleic Acids Res, 2003. **31**(16): p. 4689-95.
38. Ponnambalam, S., et al., *Transcription initiation at the Escherichia coli galactose operon promoters in the absence of the normal -35 region sequences*. J Biol Chem, 1986. **261**(34): p. 16043-8.

39. Muller-Hill, B. and P. Kolkhof, *DNA recognition and the code*. Nature, 1994. **369**(6482): p. 614.
40. Robbins, A.R., R. Guzman, and B. Rotman, *Roles of individual mgl gene products in the beta-methylgalactoside transport system of Escherichia coli K12*. J Biol Chem, 1976. **251**(10): p. 3112-6.
41. Ordal, G.W. and J. Adler, *Properties of mutants in galactose taxis and transport*. J Bacteriol, 1974. **117**(2): p. 517-26.
42. Riordan, C. and H.L. Kornberg, *Location of galP, a gene which specifies galactose permease activity, on the Escherichia coli linkage map*. Proc. R. Soc. Lond. B. Biol. Sci., 1977. **198**: p. 401 - 410.
43. Zheng, H., et al., *The prototypical H⁺/galactose symporter GalP assembles into functional trimers*. J Mol Biol, 2010. **396**(3): p. 593-601.
44. Wilson, D.B., *Source of energy for the Escherichia coli galactose transport systems induced by galactose*. J Bacteriol, 1974. **120**(2): p. 866-71.
45. Death, A. and T. Ferenci, *Between feast and famine: endogenous inducer synthesis in the adaptation of Escherichia coli to growth with limiting carbohydrates*. J Bacteriol, 1994. **176**(16): p. 5101-7.
46. Epstein, W., L.B. Rothman-Denes, and J. Hesse, *Adenosine 3':5'-cyclic monophosphate as mediator of catabolite repression in Escherichia coli*. Proc Natl Acad Sci U S A, 1975. **72**(6): p. 2300-4.
47. Hogema, B.M., et al., *Catabolite repression by glucose 6-phosphate, gluconate and lactose in Escherichia coli*. Mol Microbiol, 1997. **24**(4): p. 857-67.
48. Gorke, B. and J. Stulke, *Carbon catabolite repression in bacteria: many ways to make the most out of nutrients*. Nat Rev Microbiol, 2008. **6**(8): p. 613-24.
49. Bouffard, G.G., K.E. Rudd, and S.L. Adhya, *Dependence of lactose metabolism upon mutarotase encoded in the gal operon in Escherichia coli*. J Mol Biol, 1994. **244**(3): p. 269-78.
50. Weissborn, A.C., et al., *UTP: alpha-D-glucose-1-phosphate uridylyltransferase of Escherichia coli: isolation and DNA sequence of the galU gene and purification of the enzyme*. J Bacteriol, 1994. **176**(9): p. 2611-8.
51. Trucco, R.E., R. Caputto, and et al., *Galactokinase*. Arch Biochem, 1948. **18**(1): p. 137-46.
52. Wong, L.J. and P.A. Frey, *Galactose 1-phosphate uridylyltransferase. Isolation of a uridylyl-enzyme intermediate*. J Biol Chem, 1974. **249**(7): p. 2322-4.
53. Wilson, D.B. and D.S. Hogness, *The Enzymes of the Galactose Operon in Escherichia Coli. I. Purification and Characterization of Uridine Diphosphogalactose 4-Epimerase*. J Biol Chem, 1964. **239**: p. 2469-81.
54. Aiba, H., S. Adhya, and B. de Crombrughe, *Evidence for two functional gal promoters in intact Escherichia coli cells*. J Biol Chem, 1981. **256**(22): p. 11905-10.
55. Semsey, S., K. Virnik, and S. Adhya, *Three-stage regulation of the amphibolic gal operon: from repressosome to GalR-free DNA*. J Mol Biol, 2006. **358**(2): p. 355-63.
56. Ullmann, A., E. Joseph, and A. Danchin, *Cyclic AMP as a modulator of polarity in polycistronic transcriptional units*. Proc. Natl. Acad. Sci. U. S. A., 1979. **76**(7): p. 3194-3197.
57. Guidi-Rontani, C., A. Danchin, and A. Ullmann, *Transcriptional control of polarity in Escherichia coli by cAMP*. Mol. Gen. Genet., 1984. **195**(1-2): p. 96-100.
58. Lee, H.J., et al., *Establishment of an mRNA gradient depends on the promoter: an investigation of polarity in gene expression*. J Mol Biol, 2008. **378**(2): p. 318-27.

59. Aki, T. and S. Adhya, *Repressor induced site-specific binding of HU for transcriptional regulation*. EMBO J., 1997. **16**(12): p. 3666-3674.
60. Aki, T., H.E. Choy, and S. Adhya, *Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor*. Genes Cells, 1996. **1**(2): p. 179-188.
61. Geanacopoulos, M., et al., *Gal repressosome contains an antiparallel DNA loop*. Nat. Struct. Biol., 2001. **8**(5): p. 432-436.
62. Semsey, S., K. Virnik, and S. Adhya, *A gamut of loops: meandering DNA*. Trends Biochem Sci, 2005. **30**(6): p. 334-41.
63. Beer, M.A. and S. Tavazoie, *Predicting gene expression from sequence*. Cell, 2004. **117**(2): p. 185-98.
64. Buchler, N.E., U. Gerland, and T. Hwa, *On schemes of combinatorial transcription logic*. Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5136-41.
65. Cox, R.S., 3rd, M.G. Surette, and M.B. Elowitz, *Programming gene expression with combinatorial promoters*. Mol Syst Biol, 2007. **3**: p. 145.
66. Kauffman, S., *Homeostasis and differentiation in random genetic control networks*. Nature, 1969. **224**(215): p. 177-8.
67. Mayo, A.E., et al., *Plasticity of the cis-regulatory input function of a gene*. PLoS Biol, 2006. **4**(4): p. e45.
68. Yuh, C.H., H. Bolouri, and E.H. Davidson, *Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene*. Science, 1998. **279**(5358): p. 1896-902.
69. Hunziker, A., *Sense & Sensitivity – Bacterial Response in Good and Bad Times (PhD thesis)*. 2010 Copenhagen University, Copenhagen, Denmark.
70. Gillespie, D.T., *Exact stochastic simulation of coupled chemical reactions*. J Phys Chem, 1977. **81**: p. 2340–2361.
71. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products*. Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.
72. Geanacopoulos, M. and S. Adhya, *Functional characterization of roles of GalR and GalS as regulators of the gal regulon*. J. Bacteriol., 1997. **179**(1): p. 228-234.
73. Ryu, S., et al., *Pivotal role of amino acid at position 138 in the allosteric hinge reorientation of cAMP receptor protein*. Proc Natl Acad Sci U S A, 1993. **90**(1): p. 75-9.
74. Shin, M., et al., *DNA looping-mediated repression by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping*. Genes Dev, 2005. **19**(19): p. 2388-98.
75. Jishage, M. and A. Ishihama, *A stationary phase protein in Escherichia coli with binding activity to the major sigma subunit of RNA polymerase*. Proc Natl Acad Sci U S A, 1998. **95**(9): p. 4953-8.
76. Kolb, A., et al., *Selectivity of the Escherichia coli RNA polymerase E sigma 38 for overlapping promoters and ability to support CRP activation*. Nucleic Acids Res, 1995. **23**(5): p. 819-26.
77. Ptashne, M., *Repressors*. Current Biology, 2007. **17**(17): p. R740-741.
78. Hittinger, C.T. and S.B. Carroll, *Gene duplication and the adaptive evolution of a classic genetic switch*. Nature, 2007. **449**(7163): p. 677-81.
79. Markiewicz, P., et al., *Genetic studies of the lac repressor. XIV. Analysis of 4000 altered Escherichia coli lac repressors reveals essential and non-essential residues, as well as "spacers" which do not require a specific sequence*. J Mol Biol, 1994. **240**(5): p. 421-33.

80. Suckow, J., et al., *Genetic studies of the Lac repressor. XV: 4000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure*. J Mol Biol, 1996. **261**(4): p. 509-23.
81. Gu, Z., et al., *Duplicate genes increase gene expression diversity within and between species*. Nat Genet, 2004. **36**(6): p. 577-9.
82. Li, W.H., J. Yang, and X. Gu, *Expression divergence between duplicate genes*. Trends Genet, 2005. **21**(11): p. 602-7.
83. Shearwin, K.E., B.P. Callen, and J.B. Egan, *Transcriptional interference--a crash course*. Trends Genet, 2005. **21**(6): p. 339-45.
84. Epshtein, V., et al., *Transcription through the roadblocks: the role of RNA polymerase cooperation*. Embo J, 2003. **22**(18): p. 4719-27.
85. Courey, A.J. and S. Jia, *Transcriptional repression: the long and the short of it*. Genes Dev, 2001. **15**(21): p. 2786-96.