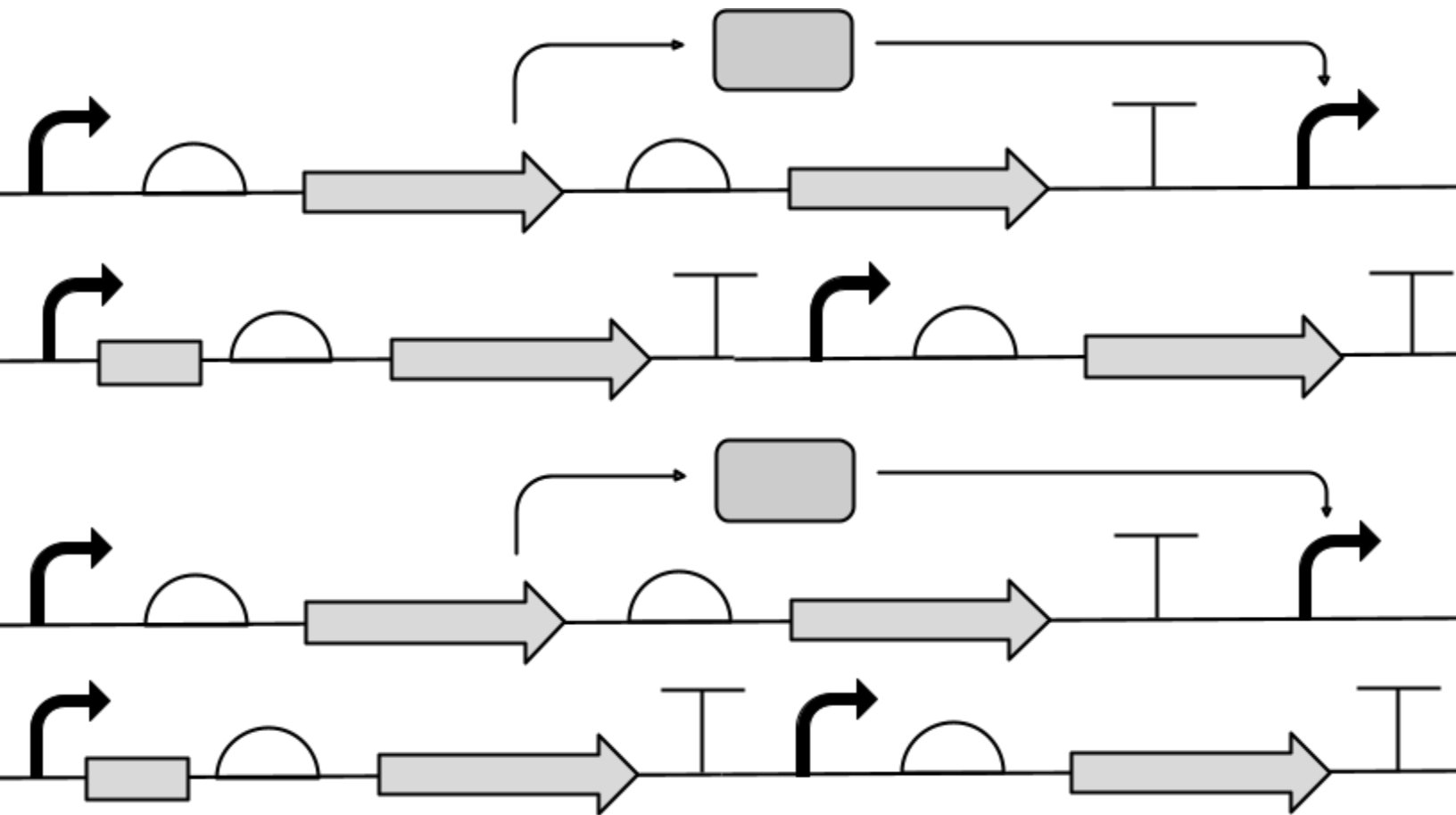


Syntetisk biologi för nybörjare 1



1. Grundläggande kunskaper

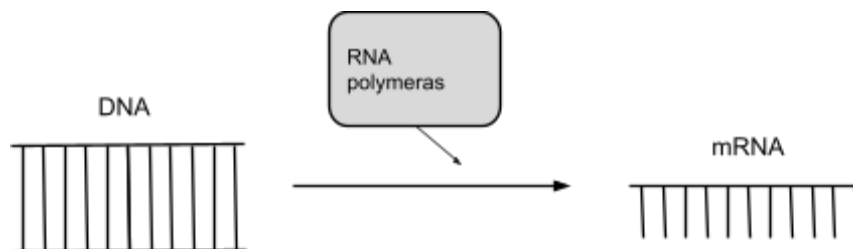
Innan vi går in på vad syntetisk biologi är och hur det funkar är det viktigt att vi benar ut vissa begrepp och koncept som är centrala för området. Många av begreppen har du säkert hört talats om innan men vissa kommer nog vara helt nya. Kom ihåg att förklaringarna jag ger inte är en fullständig representation av verkligheten och att mina förklaringsmodeller ibland faller isär om man undersöker dem tillräckligt nog. Om du är väldigt intresserad av området finns tonvis av vidare läsning online.

Inför att lära oss om syntetisk biologi är det viktigt att vi har grundläggande kunskaper kring bakterier (och andra mikroorganismer) och hur de producerar protein. För att förklara de första begreppen kan vi använda liknelsen av en kokbok. Kokboken innehåller såklart en hel drös recept, och om man följer recepten tillräckligt nog så får man en massa goda rätter och desserter. Vår kokbok är dock lite lustig. Recepten i sin ursprungliga form är nämligen fullständigt oläsbara, de är krypterade. Hursomhelst kan vi, med hjälp av ett översättningsverktyg, dechifrera det och således börja laga våra rätter.



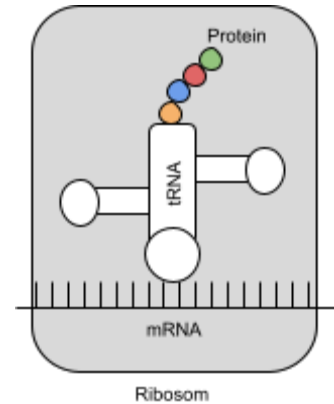
Men hur kopplar jag detta till biologi? Jo, om vi kallar kokboken för **genomet**, recepten för **gener**, dechiffreringen för **transkription**, matlagningen för **translation** och alla goda rätter för **protein** - så har vi benat ut det viktigaste konceptet i biologi: det **centrala dogmat** (dogma betyder ungefär "sanning", alltså kan det centrala dogmat översättas till biologins centrala sanning). Det centrala dogmat beskriver alltså vägen mellan **gen och protein**.

- **Transkription** är alltså dechiffreringen från ett oläsbart till ett läsbart recept. Verktöget som åstadkommer detta kallas för **RNA polymeras** (ordet polymeras kan kopplas till "att göra flera", dvs. bygga eller koppla samman). RNA polymeras gör om något oläsbart (**DNA**) till något läsbart (**mRNA**).



- **Translation** är själva matlagningen, alltså omvandlingen från den läsbara genen till ett protein. Detta sker med hjälp av ett **tRNA**. Translationen sker i små bubblor som kallas **ribosomer**.

Receptet vi pratade om ovan är även lustigt i en annan bemärkelse - det innehåller endast ingredienser bestående av 3 ord! Ingredienserna i vårt recept är proteinernas byggstenar - **aminozyror**. Det finns totalt 20 st aminozyror som bygger upp allt protein i våra kroppar och organismer. Men varför består våra ingredienser endast av 3 ord? Jo, varje aminozyra bestäms av en sekvens av 3 **nukleotider**. Nukleotider är geners byggstenar, precis som på samma sätt som aminozyror är proteinets byggstenar. De består av en socker-del, en fosfat-del och en kvävebas. Det finns 4 olika kvävebaser: **A, T, C, G (Adenin, Tymin, Cytosin och Guanin)** som utgör olika nukleotider.



Phew! Det var en krånglig kokbok! Nu kan vi förklara händelseförloppet mellan gen och protein noggrannare. Genen (som består av 4 olika nukleotider) transkriberas av ett RNA polymeras och blir nu "läsbar". Den läsbara genen translateras av ett tRNA, som bildar aminozyror av nukleotider i grupper om 3. En kedja av aminozyror bildas och omvandlas till ett protein!

Alla celler i din kropp innehåller ditt fullständiga genom, som i sig kommer från din mamma och pappa. Men varför ser alla celler olika ut? Celler på din näsa och i din mage innehåller exakt samma gener, men är fullständigt olika. Detta dilemma kan lösas med hjälp av det centrala dogmat. Alla gener blir nämligen inte till protein eftersom att celler på olika ställen i kroppen kan dechiffrera olika recept beroende på vilken miljö de är i. Olika celler har alltså olika **genuttryck**, och detsamma gäller för mikroorganismer.

Uppgifter till del 1

Använd kunskaperna från ovan och från internet för att besvara följande frågor.

1.1

Beskriv DNA, mRNA och tRNA med enkla ord eller figurer (grundläggande).

1.2

Hur många *olika* slags kombinationer om 3 finns det av nukleotider? Hur många aminosyror finns det?

Om det är olika många, hur kan denna skillnad förklaras? (Svaret är inte så komplicerat som man kan tro!)

1.3

Beskriv det centrala dogmat med enkla ord eller figurer.

1.4

Använd internet eller din förkunskap för att ta reda på var DNA, mRNA och tRNA finns, samt var transkription och translation sker. Rita gärna en figur.

Facit till del 1

1.1. DNA är en dubbelsträngad helix som består av basparen A,C,T och G. Till skillnad från DNA är RNA enkelsträngat och basparet T byts ut mot U (uracil). mRNA är det RNA som translateras till protein. tRNA har en annorlunda form (korsliknande), och är det RNA som används under translationen.

1.2. Det finns 4 olika nukleotider, och 3 nukleotider kodar för 1 aminosyra. Därför finns det $4^3=64$ olika kombinationer av nukleotidsekvenser, men det finns bara 20 olika aminosyror. Skillnaden förklaras då det finns flera nukleotidsekvenser som kodar för samma aminosyra!

1.3 Det centrala dogmat beskriver vägen från gen till protein. DNA:t omskrivs till mRNA i transkriptionen (dechiffrering). Därefter translateras mRNA:t till aminosyror med hjälp av tRNA, och proteinet skapas.

1.4 DNA finns i cellkärnan. mRNA transkriberas i cellkärnan men transporteras sedan till det endoplasmatiska retiklet. tRNA finns i ribosomer. Translationen sker på ytan av det endoplasmatiska retiklet.

Begreppslista till del 1

Genomet - Den fullständiga genuppsättningen hos en organism. Alltså hela kokboken!

Gen - Ett segment i genomet. Ett recept i kokboken.

Transkription - Omskrivningen från DNA till RNA. Dechiffrering av kokboken.

Translation - Omskrivningen från mRNA till aminosyrasekvens med hjälp av tRNA. Matlagning!

Centrala dogmat - Vägen från gen till protein.

DNA - Dubbelstrandad polynukleotid. Innehåller arvsmassan och finns i cellkärnan.

RNA - Enkelstrandad polynukleotid.

- **mRNA** - RNA som skapas via transkription och som kodar för protein. Kommer translateras.
- **tRNA** - Komplex RNA-struktur som gör om mRNA till en aminosyrasekvens genom translationen.

RNA polymeras - Enzymet (proteinet) som bygger RNA vid transkription.

Ribosom - Struktur där translationen sker.

Aminosyror - Byggstenar till proteiner.

Nukleotider - Byggstenar till DNA och RNA.

Genuttryck - Hur mycket av en gen som transkriberas och uttrycks.

2. Syntetisk biologi

Syntetisk (konstgjord, artificiell) biologi är ett gryende område inom bioteknik. Området går ut på att förändra recept i organismers kokbok och på så sätt få dem att baka rätter som dom annars inte skulle få för sig att baka, så att vi kan dra nytta av dem på nya sätt - alltså ändrar vi organismers genom för att få dem att producera nya protein. Ofta modifieras genomet hos bakterier eller jäst för att åstadkomma saker som tidigare var omöjligt för organismen, och som vi nu kan nyttja.

Inom fältet är man överens om att informationen ska vara så tydlig och tillgänglig som möjligt. Man ska inte behöva vara speciellt insatt för att greppa de grundläggande koncepten. Därför finns ett **register**. Registret liknar en lego-katalog, men istället för legobitar finns det något som kallas **BioBricks** (gen-bitar). I registret finns även information om gen-biten, hur den fungerar och hur den kan sättas ihop med andra gen-bitar.

För att särskilja på gener och proteiner använder vi oftast *kursivt* för att denotera gener, och vanligt text samt stor bokstav för att denotera protein. Exempelvis så kodar *erik* för proteinet Erik.

2.1 BioBricks (gen-bitar)

En av de viktigaste bitarna kallas **promotor**. Vid promotorn startar transkriptionen genom att RNA polymeraset fäster till den. Genom att byta promotor kan man styra genuttrycket, och således hur mycket protein som skapas. Det finns två typer av promotorer: **konstitutiva** och **inducerbara**. En konstitutiv promotor tillåter RNA polymeraset att fästa när den vill, och genuttrycket är därför kontinuerligt. Inducerbara promotorer är promotorer som endast låter RNA polymeraset fästa då speciella ämnen är närvarande, och kan användas för att styra genuttrycket. Promotorn denoteras med böjda pilar.

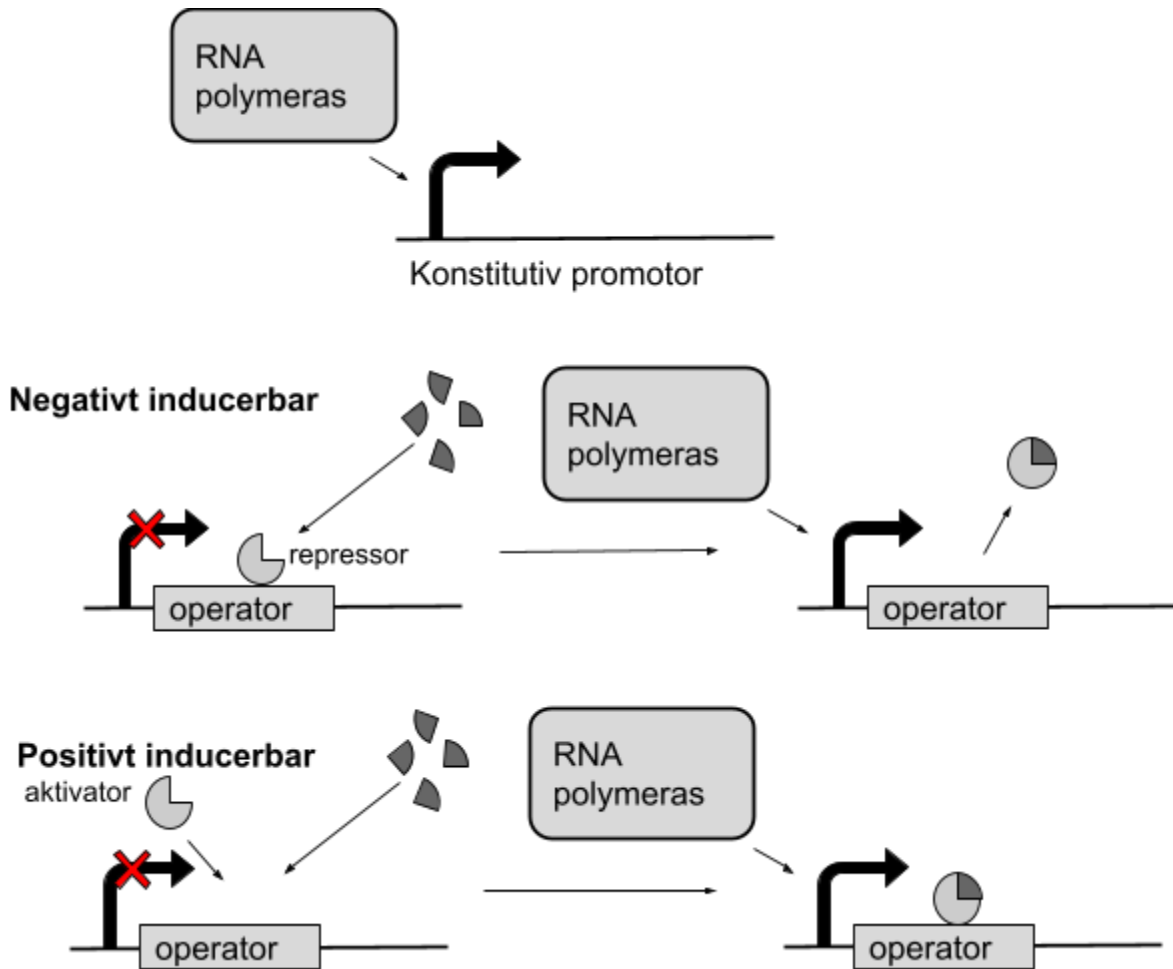
Innan vi pratar om de olika typerna av inducerbara promotorer är det viktigt att vi nämner vad en **operator** är. En operator är en gen-bit som sitter intill en inducerbar promotor, och där ämnen kan fästa för att styra induktionen (uttrycket).

Det finns två sorters inducerbara promotorer: **negativt inducerbara** och **positivt inducerbara**. En negativt inducerbar promotor styrs av en **repressor**. Repressorn sitter på operatoren och hämmar transkribering, och släpper från operatoren då en substans är närvarande. Alltså uttrycks endast generna som följer efter den inducerbara promotorn då substansen är närvarande.

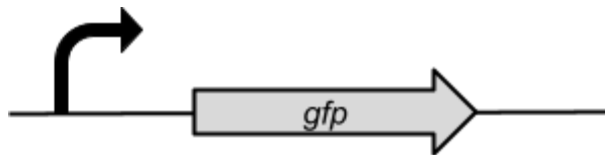
En positivt inducerbar promotor kräver att en **aktivator** fäster till operatoren för att transkription ska ske. En aktivator är ett protein som som fäster på operatoren, ofta om ett specifikt substrat är fäst till den. Om substratet *inte* fäster till aktivatorn fäster aktivatorn inte till operatoren, och RNA

polymeraset kan inte börja transkribera. Således sker transkription **endast** om substratet förekommer, och fäster till aktivatorn. (OBS! Det finns även aktivatorer som inte kräver ett visst substrat, utan fäster till operatoren oavsett!)

Alltså: **negativt inducerbar** - repressorn släpper för att RNA polymeraset ska fästa, **positivt inducerbar** - aktivator fäster för att RNA polymeraset ska fästa.



Vid promotorn börjar alltså transkriptionen. Efter promotorn sitter **proteinkodande gener** som ska uttryckas. Exempelvis kan man sätta en gen för ett grönljysande protein, green fluorescent protein (GFP), efter en promotor. Då börjar GFP produceras och cellen börjar lysa grönt! Proteinkodande gener denoteras med en stor, rak pil, innehållande genens namn.



Nästa bit krävs för att translationen ska ske, genom att erbjuda ribosomen en plats att sitta på. Biten kallas för **ribosome-binding-site (RBS)** och denoteras som en halvcirkel. Om man har många proteinkodande gener krävs det ett RBS inför varje bit, annars fäster inte ribosomen, och translationen kan inte ske.



För att transkriptionen ska sluta krävs något som kallas en **terminator**. Terminatoren består av sekvenser som signalerar till RNA polymeraset att det är dags att hoppa av! Terminatoren denoteras som ett T.



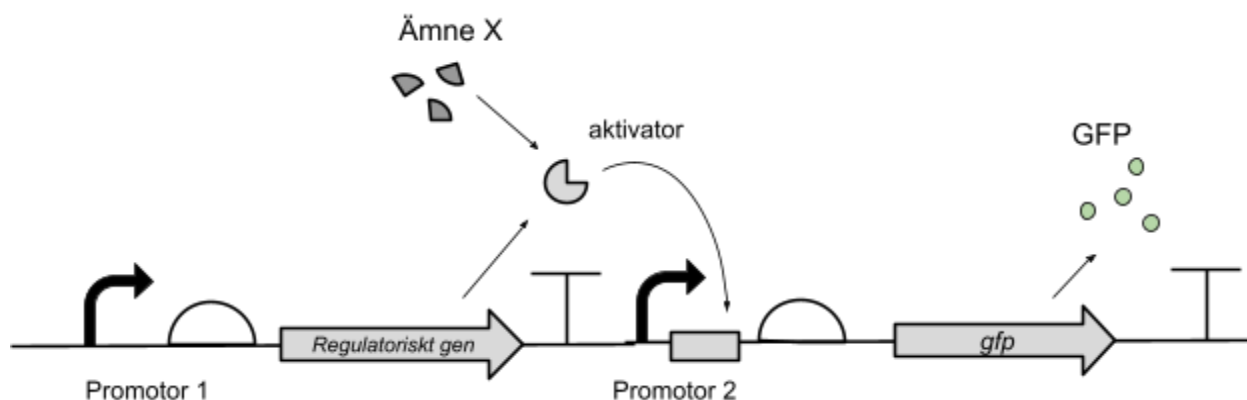
Platsen mellan promotor och terminator innehåller ofta besläktade gener och kallas **operon**. Eftersom alla gener i operonet styrs av samma promotor blir deras genuttryck samma.

2.2 Proteiner och gener kan samspela!

Nu kan vi alla de viktigaste delarna: promotor, proteinkodande gen, RBS, och terminator. En stor del av arbetet inom syntetisk biologi går ut på att undersöka hur proteiner och gener samspelar. På så sätt kan man skapa stora gensekvenser som utför fantastiskt komplicerade saker.

En viktig interaktion mellan protein och gen är något vi redan utforskat: den inducerbara promotorn. Inducerbara promotorer styrs av en aktivator eller en repressor, och jag nämnde att dessa ofta var proteiner. Proteiner som fäster till en operator och styr uttrycket av operonet kallas **regulatoriska proteiner**. Ofta betecknas regulatoriska proteiner med ett 'R' samt ett namn som hänvisar till operonet det fäster vid. Exempelvis fäster det regulatoriska proteinet ErikR till promotorn i erik-operonet.

Nedan är ett exempel på en hypotetisk sensor som kan skapas av de gen-bitar vi lärt oss. Sensorn detekterar ämne X och producerar GFP i dess närvaro.



Men hur fungerar sensorn? Jo, för det första kan vi se att promotor 1 är en konstitutiv promotor, och den följande regulatoriska genen uttrycks och translateras till ett regulatoriskt protein. Det regulatoriska proteinet är en aktivator som kräver ämne X för att fästa till operatoren vid promotor 2. Promotor 2 är alltså en positivt inducerbar promotor. Då ämne X förekommer kommer aktivatorn att fästa till operatoren, alltså sker transkribering, och GFP uttrycks!

Uppgifter till del 2

Använd kunskaperna från ovan och från internet för att besvara följande frågor.

2.1

Beskriv kortfattat begreppen inducerbar och konstitutiv promotor, proteinkodande gener, RBS, terminator och operon. Rita gärna figurer!

2.2

Varför sätts promotorn i början av sekvensen, och varför sätts terminatorn i slutet?

2.3

Varför behövs inget RBS innan promotorn eller terminatorn?

2.4

Om jag vill ha maximalt genuttryck för en gen, ska jag använda en inducerbar eller konstitutiv promotor framför den?

Facit till del 2

2.1. Promotorn sitter i början av gensekvenser och styr genuttrycket då RNA polymeraset fäster vid den för att starta transkriptionen. Konstitutiva promotorer tillåter alltid RNA polymeraset att fästa. Inducerbara promotorer har en operator-region, där aktivatorer/repressorer kan fästa för att styra genuttrycket.

Proteinkodande gener är gener som resulterar i ett protein efter translation.

RBS (ribosome binding site) är en plats som ribosomen känner igen och fäster vid. Detta är viktigt vid translation.

Terminatorn sätts i slutet av sekvensen och avslutar transkriptionen genom att få RNA polymeraset att hoppa av.

Ett operon innehåller alla gener från promotor till terminator.

2.2 Promotorn sitter i början av sekvensen eftersom den initierar transkriptionen. Om den suttit i mitten hade de proteinkodande generna innan inte blivit transkriberade!

Terminatorn sitter i slutet eftersom att den avslutar transkriptionen. Om den suttit i mitten hade de proteinkodande generna efter inte blivit transkriberade!

2.3 Eftersom att promotorn och terminatorn inte ska translateras till protein krävs inget RBS.

2.4 För att få maximalt genuttryck används oftast en konstitutiv promotor! Detta då den alltid tillåter RNA polymeraset att fästa.

Begreppslista del 2

BioBrick-register - Registret som innehåller information om gen-bitar.

BioBrick - En gen-bit med specificerad funktion i registret.

Promotor - Gen som sitter först i operonet, och som styr genuttrycket.

- **Konstitutiv promotor** - Promotor där RNA polymeras alltid kan fästa.
- **Inducerbar promotor** - Promotor där särskilda omständigheter krävs för att RNA polymeraset ska fästa.
 - **Positivt inducerbar** - Promotor som kräver en aktivator, som sedan fäster till operatoren, för att RNA polymeraset ska fästa.
 - **Negativt inducerbar** - Promotor som kräver att en repressor släpper från operatoren för att RNA polymeraset ska fästa.

Operator - Gen-bit som är belägen nära en inducerbar promotor och som bestämmer huruvida ett RNA polymeras ska fästa.

Aktivator - Protein som fäster till operatoren så att RNA polymeraset kan fästa vid en positivt inducerbar promotor. Aktivatorn kräver ofta ett visst ämne för att fästa till operatoren.

Repressor - Protein som släpper från operatoren så att RNA polymeraset kan fästa vid en negativt inducerbar promotor. Repressorn kräver ofta ett visst ämne för att släppa från operatoren.

Proteinkodnande gen - Gen-bit som kommer transkriberas och translateras till ett protein.

Ribosome-binding-site (RBS) - Ställe för ribosomen att fästa på, vilket krävs för att translation ska ske.

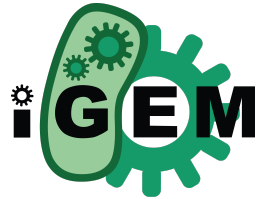
Terminator - Gen-bit som säger till RNA polymeraset att släppa och därmed avsluta transkriptionen.

Operon - Bit i genomet som innehåller gener som är utformade för ett enat syfte, och som omsluts av en promotor och terminator.

Regulatoriska protein - Protein som fäster till operatoren och därmed styr genexpressionen.

3. Att använda registret

Registret som vi pratade om i del 2 ligger under en organisation som heter iGEM. iGEM står för international Genetically Engineered Machine. Årligen tävlar olika universitet och skolor från världen över med projekt i syntetisk biologi i den stora iGEM-tävlingen. Därför kan du ofta se iGEMs logga om du stöter på projekt eller information kring syntetisk biologi.



Att använda registret kan vara lite knepigt, speciellt inte om man är van vid vokabuläret. Hursomhelst kan det vara bra att öva på, genom att gå in och titta. Registret hittas om du följer länken: http://parts.igem.org/Main_Page.

Varje BioBrick har en kod som gör att man kan hitta den i registret. Samtliga koder börjar med BBa_ följt av en unik kod. Exempelvis har GFP koden BBa_E0040.

Uppgifter till del 3

3.1 Hitta GFP i registret och undersök vilket djur proteinet kommer ifrån.

3.2

Vilken egenskap har proteinet med koden BBa_K3282002?

3.3

Kan du hitta någon likhet mellan GFP och BBa_J04450?

3.4

Är BBa_R0010 en konstitutiv eller en inducerbar promotor?

3.5

Nu är det dags för dig att hitta på en egen gensekvens. Använd bitar du hittar i registret, eller bitar som du själv hittar på. Tänk på de grundläggande principer du lärt dig i detta kompendium! Rita ut din sekvens på ett korrekt sätt och förklara vad den gör! Visa den för Erik och berätta.

Facit till del 3

3.1 GFP kommer från maneten *Aequorea victoria*.

3.2 Det binder till bly.

3.3 BBa_J04450 kodar för RFP (red fluorescent protein) och lyser, liksom GFP.

3.4 BBa_R0010 är en negativt inducerbar promotor, vars induktion huvudsakligen styrs av proteinet Lacl.

Feedback

Ge mig gärna er feedback på detta material och era tankar om området! Allt uppskattas!
