

# Kleine genetische Netze Einfache Steuerungsprogramme am Beispiel der SOS DNS Reparatur in E. coli

Marvin Schulz

21. Februar 2006

Der Zelle stehen eine ganze Reihe von Mechanismen zur Verfügung um ihren Stoffwechsel zu regulieren. Einer von ihnen ist die Regulation der Transkription, welche effektiv die Menge an vorhandenem Protein und damit auch die Stoffwechselrate kontrolliert. Die Steuerung der Transkriptionsrate erfolgt entweder global für alle Gene, über die Menge an vorhandener RNA-Polymerase und ihr assoziierten Cofaktoren, oder lokal für eine kleine Gruppe von Genen über sogenannte Transkriptionsfaktoren.

Im Folgenden beschäftigen wir uns nun mit den Möglichkeiten, die der Zelle zur Verfügung stehen, wenn sie ein kleines genetisches Netzwerk mit nur einem einzigen Transkriptionsfaktor steuert. Dazu betrachten wir das SOS DNS Reparatursystem in E. coli mit seinen kinetischen Parametern, die experimentell geschätzt wurden. Das mathematische Modell für dieses System wurde von Michal Ronen et.al. im Artikel *Assigning numbers to the arrows: Parameterizing a gene regulation network by using accurate expression kinetics* vorgeschlagen.

## 1 Vereinfachende Annahmen

Nach dem heutigen Kenntnisstand ist die Umsetzung von genetischen Informationen in Proteine ein komplizierter Prozess, der eine ganze Reihe von Schritten umfasst. Da die Integration all dieser Schritte in unser Modell einen großen Aufwand bedeuten würde, treffen wir ein paar vereinfachende Annahmen.

- Die effektive Proteinmenge hängt nur von der Rate der Transkription seiner RNS und seiner Abbaurate ab.  
Hiermit vernachlässigen wir jegliche Einflüsse aus posttranskriptionellen Modifikationen und der Translation.
- Die effektive Transkriptionsrate hängt direkt von der Menge des gebundenen Transkriptionsfaktors ab.  
Diese Annahme ignoriert jegliche Einflüsse von anderen in die Transkription involvierten Proteinen.
- Der Transkriptionsfaktor wirkt nicht kooperativ.  
Dies bedeutet für uns, dass wir eine entscheidende Vereinfachung der Bindungskinetik des Transkriptionsfaktors an die DNS haben, da wir Effekte

vernachlässigen, die auftreten, wenn sich zuerst mehrere dieser Proteine zusammenlagern müssen, bevor sie an die DNS binden können.

Somit modellieren wir nur die Bindung eines Transkriptionsfaktors an die DNS und nehmen an, dass die Menge an gebundenem Faktor gleichbedeutend ist mit der Produktionsrate des von uns betrachteten Proteins.

## 2 Modellierung der Transkriptionsrate

Zur Beschreibung des Bindungsverhaltens wählen wir die Michaelis-Menten-Kinetik. Diese wurde ursprünglich entwickelt um irreversible Reaktionen mit nur einem Substrat und einem Enzym zu beschreiben. Dabei geht man von den folgenden Annahmen aus.

Erstens, dass sich Enzym und Substrat relativ häufig zu einem Komplex verbinden und sich auch wieder trennen, eine Umsetzung vom Substrat zum Produkt jedoch eher selten ist, zweitens, dass ein stationärer Zustand existiert, in dem sich die effektive Konzentration an Enzym-Substrat-Komplex nicht ändert und drittens, dass das Enzym in wesentlich geringeren Konzentrationen vorliegt als das Substrat. Unter all diesen Annahmen, ergibt sich die folgende Gleichung für die Reaktionsgeschwindigkeit:  $v = \frac{V_{max} \cdot S}{S + K_m}$

Diese Annahmen sind insofern erfüllt, als dass erstens die Transkription im Gegensatz zur Bindung des Faktors an die DNS ein eher seltenes Ereignis darstellt, zweitens die Konzentration von DNS-Faktor-Komplex einen stationären Zustand annehmen dürfte und drittens mehr Transkriptionsfaktormoleküle als Bindestellen vorhanden sind.

Da allerdings bei der Michaelis-Menten-Kinetik, angewandt auf unser Modell, die Produktion der RNS mit zunehmender Menge des Transkriptionsfaktors steigen würde, müssen wir zur Modellierung des repressiven Effektes des Faktors die Kinetik geringfügig modifizieren. Hierbei ist anzunehmen, dass in Abwesenheit des Repressors die Transkriptionsrate maximal ist und sich proportional zur Menge des gebundenen Repressors reduziert.

Aus dieser Herleitung ergibt sich die Form:  $v = V_{max} - \frac{V_{max} \cdot S}{K_m + S} = \frac{V_{max} \cdot K_m}{K_m + S}$   
Die Dynamik der Transkription bei einem aktivierenden oder einem reprimierenden Faktor ist in Abbildung 1 dargestellt. Wie man in der Darstellung sieht, ist auch bei der Repressorbindungskinetik  $K_m$  die charakteristische Konzentration, bei der die Transkriptionsrate auf die Hälfte ihres Maximalwertes verringert ist. Die Größe des  $K_m$  Wertes wirkt sich insofern auf die Transkriptionsrate aus, als dass bei einer Kinetik mit hohem  $K_m$  mehr Faktor vorhanden sein muss, damit ein repressiver Effekt auftritt.  $K_m$  beschreibt also die Affinität des Faktors für seine Bindungsstelle auf der DNS.

Simuliert man nun einmal Veränderungen der Konzentration des Transkriptionsfaktors, so sieht man, dass sich das Gen mit der weniger affinen Bindungsstelle bei einem Abfall der Faktorkonzentration schneller von der Repression erholt und sein Protein in der Zelle zuerst akkumuliert. Dieses Gen wird dann auch bei einem erneuten Anstieg der Konzentration des Transkriptionsfaktors erst später wieder abgeschaltet als ein Gen mit stärker affiner Bindungsstelle und seine Produkte bleiben wesentlich länger in der Zelle vorhanden.

Dieses Verhalten macht sich die Zelle zu nutze, um mit nur einem Transkriptionsfaktor eine Art Phasenverhalten in seiner metabolischen Antwort zu

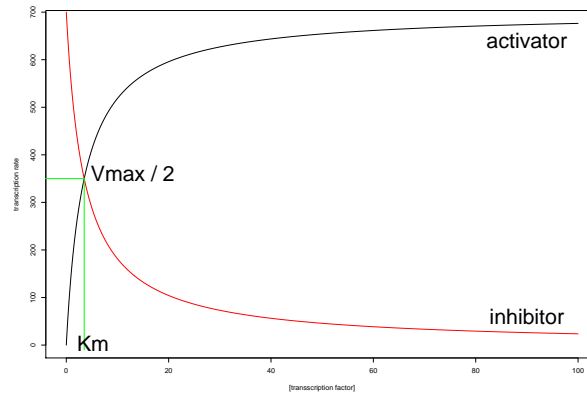


Abbildung 1: Relative Transkriptionsrate in Abhängigkeit der relativen Konzentration des Aktivators oder Repressors

erzeugen. Ein einfaches Beispiel hierfür ist die SOS DNS Reparatur in *E. coli.*, für das Ronen et. al. ein parametrisiertes Modell entworfen haben.

### 3 Topologie des Modells

Die SOS DNS Reparatur ist ein System, welches es *E. coli* ermöglicht, sehr schnell Schäden an seiner DNS zu beheben. Diese Art der Reparatur ist allerdings auch sehr fehleranfällig, weshalb sie meist für das Individuum letal ist. Evolutionär ist es trotzdem sinnvoll, da einige wenige Bakterien diesen Vorgang überleben. Gesteuert wird das System über einen Transkriptionsfaktor LexA, welcher alle Gene in diesem System (unter anderem auch sein eigenes) reprimiert und ein *Sensorprotein* RecA, welches in der Lage ist, Schäden in der DNS zu erkennen und daraufhin die Selbstspaltung des LexA anregt.

Sollten nun aus irgendeinem Grund, z.B. durch ionisierende Strahlung, Schäden in der DNS auftreten, über den die DNS-Polymerase nicht hinweglesen kann, so wird durch die fortschreitende Entwindung der Doppelhelix ein Abschnitt einzelsträngiger Nukleinsäure frei. In vivo werden Einzelstränge durch das ss-DNS bindende Protein abgeschirmt, in der SOS Antwort hingegen vermitteln bestimmte Proteine die Bindung von RecA an die DNS. Diese Bindung aktiviert RecA, welches jetzt LexA zur Selbstspaltung bringen kann. Dadurch werden eine ganze Reihe von Genen schwächer reprimiert und die Transkription an ihnen nimmt zu. Ihre Produkte beheben im Folgenden verschiedene Schäden an der DNS, so dass die Replikation fortfahren kann. Sinkt nun die Konzentration an einzelsträngiger DNS, wird ebenfalls das aktivierte RecA weniger, was dazu führt, dass LexA nicht mehr gespalten wird und in der Zelle akkumuliert, bis es sich wieder in seinem alten Gleichgewicht befindet.

Innerhalb dieses Gennetzwerkes befinden sich eine ganze Reihe an verschiedensten Proteinen, denen wir im Folgenden besondere Aufmerksamkeit widmen:

**uvrA**, welche ein Teil der ABC Exonuklease ist, die für die Entfernung einzelner falscher Nukleotide aus der DNS verantwortlich ist,

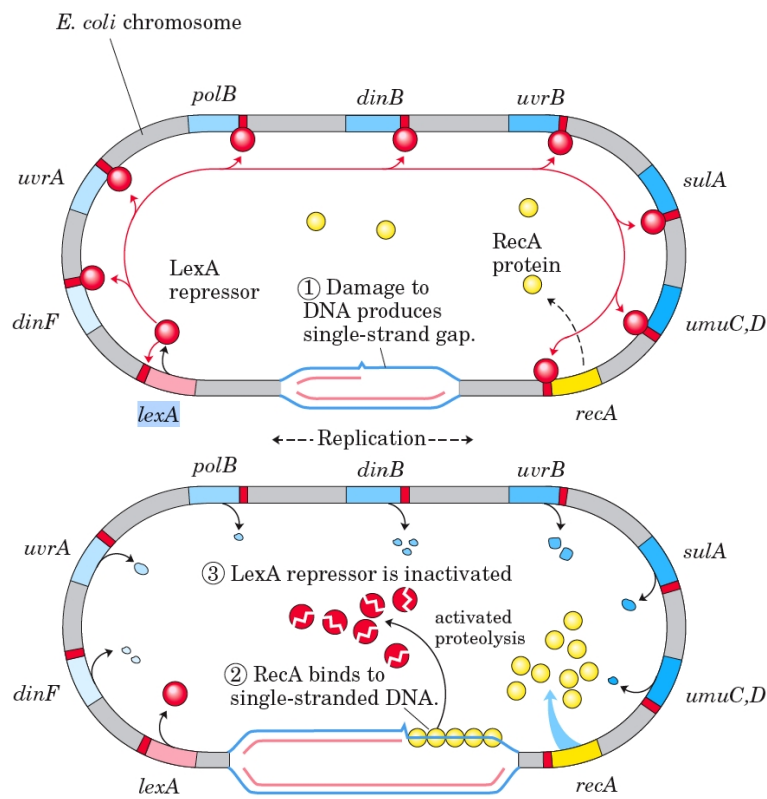


Abbildung 2: Übersicht des SOS DNS Reparatursystems, Quelle: Lehninger Biochemie; Nelson & Cox; 4. Auflage

**umuD** , welche ein Teil der replikationsinitialisierenden DNS-Polymerase V ist und

**polB (dinA)** , ein Teil der DNS-Polymerase II, welche zur Elongation der Replikation gebraucht wird.

Eine detaillierte Liste aller im Netzwerk aktivierten Gene findet sich im Anhang.

Somit enthält diese Liste Proteine, die an einer frühen, mittleren und späten metabolischen Antwort auf Schäden in der DNS beteiligt sind. Nun wäre es für die Zelle denkbar unsinnig alle diese Enzyme für die ganze Zeit des Reparaturprozesses in hohen Konzentrationen bereit zu halten, da sie nicht immer benötigt werden. Die nötige Steuerung des Systems übernimmt LexA, wobei die unterschiedlichen Expressionsprofile der einzelnen Gene, wie oben bereits besprochen, über unterschiedliche Affinitäten des Transkriptionsfaktors an die entsprechenden Bindungsstellen erzeugt werden. Werte für diese unterschiedlichen Affinitäten wurden von Ronen et. al. experimentell bestimmt.

## 4 Bestimmung der Parameter

Um für alle Gene in Echtzeit ihre Transkriptionsrate zu bestimmen, brachten die Experimentatoren Reporterplasmide in *E. coli* Zellen ein. Diese enthielten jeweils einen Operator mit der DNS Bindungsstelle des zu untersuchenden Gens und ein dahinterkloniertes Gen für das Grün Fluoreszierende Protein (GFP). Auf Grund der geringen Abbaurrate von GFP konnte im Experiment die Zunahme der Fluoreszenz innerhalb der Zelle mit der Transkriptionsrate hinter unserem untersuchten Operator gleich gesetzt werden. Diese Fluoreszenzmessungen wurden nun für sämtliche Zellen nach einer Bestrahlung mit UV-Licht, welche den Zweck hatte, Schäden in der DNS hervorzurufen, durchgeführt.

Die durch die Differenzierung der Fluoreszenzmessungskurve gewonnenen Zeitverläufe der transkriptionellen Aktivität wurden zuerst durch einen Gauschen Median Filter und einen Fit an eine polynomiale Funktion geglättet und danach durch eine Regression an unsere oben genannte Funktion angepasst.

Die durch die Regression von Ronen et. al. gewonnenen Werte sind hier aufgeführt:

<i>Gen</i>	$K_m$	$V_{max}$
<u>uvrA</u>	0.09 ±0.04	2800 ±300
<u>lexA</u>	0.15 ±0.08	2200 ±100
<u>recA</u>	0.16 ±0.07	3300 ±200
<u>umuD</u>	0.19 ±0.1	330 ±30
<u>polB</u>	0.35 ±0.15	70 ±10
<u>ruvA</u>	0.37 ±0.1	30 ±2
<u>uvrD</u>	0.65 ±0.3	170 ±20
<u>uvrY</u>	0.51 ±0.25	300 ±200

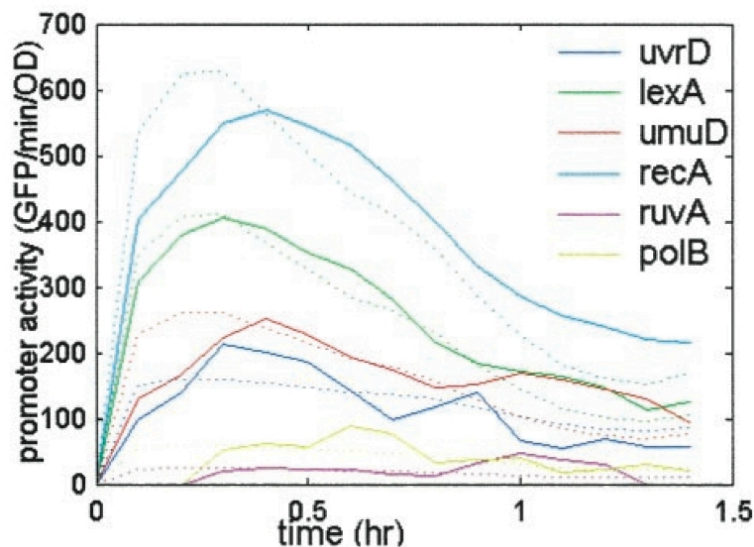


Abbildung 3: Gemessene (durchgezogene) und vorhergesagte (gepunktete Linie) Promotoraktivität, Quelle: Assigning numbers to the arrows, Ronen et. al.

Für einige dieser Gene wie z.B. polB weichen die vorhergesagten Werte sehr

von den gemessenen ab. Dies deutet darauf hin, dass einige Gene unseres Systems weiteren Regulationsmechanismen unterliegen, die wir nicht in unser Modell aufgenommen haben.

Es wurde ebenfalls überprüft, ob die Regression auf eine Hill Kinetik ( $V = \frac{V_{max} S^k}{S^k + S_{0,5}^k}$ ) deutlich bessere Ergebnisse liefert. Da hier allerdings die Regression für den Hill-Koeffizient  $k$  auf einen Wert von etwa 1 kam, was keinerlei Kooperativität zwischen den Transkriptionsfaktoren entspricht, wurde diese Kinetik nicht näher betrachtet.

## 5 Zeitliches Verhalten des Netzwerkes

Nun, da wir die mathematischen Parameter unseres Systems wissen, können wir uns Gedanken über das zeitliche Verhalten des Gesamtnetzwerkes machen.

Wie wir bereits wissen, sorgte ein großer  $K_m$ -Wert bei der Transkriptionsfaktorbindungskinetik dafür, dass der Faktor auch in relativ hohen Konzentrationen nur wenig an seine Bindungsstelle gebunden ist. Somit haben wir mit  $polB$  ein Gen, welches bereits früh bei einem Abfall der Konzentration von LexA exprimiert wird und erst spät in der Akkumulationsphase von LexA wieder reprimiert wird. Die Expression von  $uvrA$  hingegen nimmt erst bei wesentlich geringeren Repressorkonzentrationen zu und auch bereits bei geringeren Konzentrationen wieder ab.

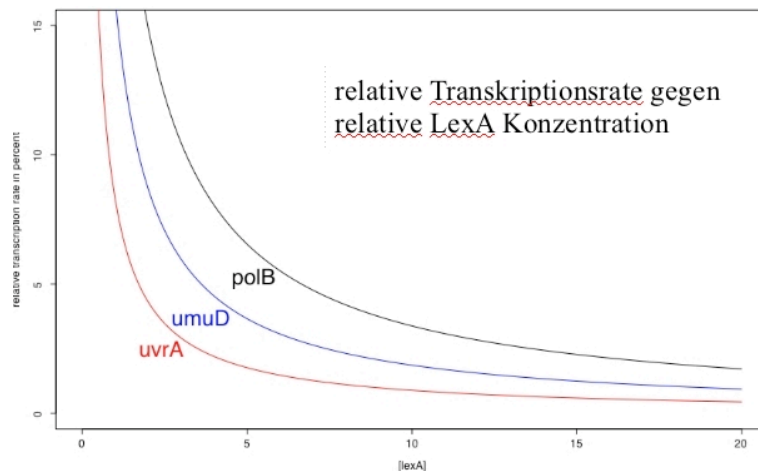


Abbildung 4: relative Transkriptionsrate in Abhängigkeit der relativen LexA Konzentration

Im Zusammenhang mit der Dynamik des LexA Ab- und Aufbaus macht dies aber wiederum Sinn. Wie man in Abbildung 5 sieht, wird LexA anfangs relativ schnell abgebaut, während seine erneute Akkumulierung deutlich mehr Zeit in Anspruch nimmt. Somit werden am Anfang alle Gene trotz der unterschiedlichen Affinitäten des Transkriptionsfaktors für sie fast gleichzeitig aktiviert. Die frühzeitigere Expression von  $polB$  bringt also für die Zelle keine gravierenden Nachteile, zumal die Transkriptionsgeschwindigkeit dieses Gens nicht sonderlich hoch ist. Der langsamere Aufbau von LexA hingegen wirkt sich schon deutlicher

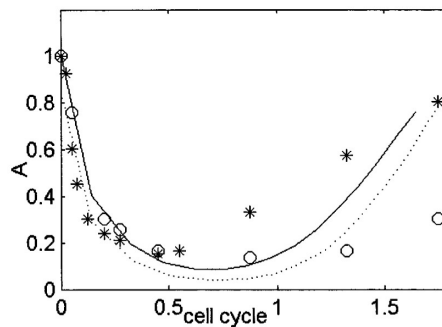


Abbildung 5: relative Transkriptionsrate in Abhängigkeit der relativen LexA Konzentration; Quelle: Assigning numbers to the arrows, Ronen et. al.

auf die Expressionsdauer der einzelnen Gene aus. Es wird also die Produktion des in der Frühphase wichtigen UvrA wesentlich früher gestoppt als die von PolB, welches im späteren Reparaturprozess eine Rolle spielt. Genauso sieht es auch mit allen anderen Genen in diesem System aus, je früher ihre Produkte im Reparaturprozess benötigt werden, desto geringer sind die  $K_m$ -Werte ihrer Transkriptionsfaktorbindungskinetiken. Dabei liegen zwischen den Inaktivierungen der einzelnen Gene jeweils etwa 10 Minuten.

## 6 Ausblick

Generell lassen sich für Gennetzwerke dieser Art folgende Möglichkeiten der Steuerung ableiten. Die Aktivierungs- und Deaktivierungsreihenfolge der Expression an den einzelnen Genen wird gesteuert über die Affinität des Transkriptionsfaktors für seine entsprechende Bindungsstelle. Dies entspricht dabei dem  $K_m$  Wert der Bindungskinetik. Die Intervalle zwischen den Aktivierungen und Inaktivierungen der unterschiedlichen Gene kann über die Geschwindigkeit des Auf- und Abbaus des Transkriptionsfaktors gesteuert werden. Eine langsame Veränderung seiner Konzentration entspricht dabei längeren Intervallen zwischen den Aktivitätsveränderungen der jeweiligen Gene. Somit stehen hier drei Möglichkeiten für Genaktivierungsmuster zur Verfügung: erstens die differentielle Aktivierung und gleichzeitige Inaktivierung von Genen, zweitens eine fast simultane Aktivierung und stufenweise Inaktivierung und drittens sowohl gestaffelte Aktivierung als auch Inaktivierung.

Selbst für Systeme die nur von einem Transkriptionsfaktor gesteuert werden, ergeben sich also schon diverse komplexe Steuerungsmöglichkeiten

Die Erkenntnisse, die dieser Artikel gebracht hat, lassen sich auch noch in anderer Art verwenden. Durch die Werte die wir für die einzelnen Kinetiken bekommen haben, sind wir in der Lage die Expression der jeweiligen Gene nur aus der Konzentration von LexA vorherzusagen. Wenn wir nun auch noch den Abbau der Enzyme beschreiben, ist es uns möglich, die einzelnen Konzentrationen der Proteine unseres Netzwerkes nur durch den Konzentrationsverlauf von LexA zu beschreiben. Somit können künftige Experimente darauf verzichten, alle Enzyme des Netzwerkes in ihrer Konzentration zu messen.

Es sollte ebenfalls möglich sein,  $K_m$  Werte für die Bindung des Transkriptionsfaktors aus Alignments der zu untersuchenden DNS Bindungsstelle mit einer Konsensussequenz bekannter Regulons abzuleiten. Letztere kann aus multiplen Alignments von Bindungsstellen für die der Transkriptionsfaktor eine hohe Affinität besitzt, gewonnen werden. Für LexA wurde diese Konsensussequenz von Fernandez De Henestrosa et. al. ermittelt: (5'-TACTG(TA)5CAGTA-3'). Über sie kann man nun Aussagen über die Affinität anderer Regulons für ihren Transkriptionsfaktor machen.

## A Proteine im Netzwerk

Diese Liste an Proteinen enthält 13 der geschätzten 70 Gene, welche von LexA reguliert werden. Bei den anderen ist entweder die Rolle in der DNS Reparatur noch ungeklärt oder das codierte Protein unbekannt.

**polB (dinA)** Teil der DNS Polymerase II, wird in der Elongation der Replikation gebraucht

**uvrA & uvrB** Teile der ABC Exonuklease zur Entfernung einzelner falsch paarender Nukleotide aus der DNS

**umuC & umuD** Bilden die DNS Polymerase V, wird in der Initiation der Replikation gebraucht

**sulA** Codiert für ein Protein, welches in der Lage ist, den Zellzyklus anzuhalten

**recA** Sensor für DNS Schäden

**dinB** Codiert für die DNS Polymerase IV

**lexA** Transkriptionsrepressor des SOS DNS Reparatursystems

**ssb** Codiert für das Einzelstrang bindende Protein

**uvrD** Codiert für die DNS Helikase II, welche doppelsträngige DNS entwindet

**himA** Teil des Integration Host Factors

**recN** Codiert für ein Protein, welches in der Rekombinationsreparatur benötigt wird

## B Quellen

- Ronen et.al.: Assigning numbers to the arrows: Parameterizing a gene regulation network by using accurate expression kinetics
- <http://de.wikipedia.org/wiki/SOS-Antwort>
- Nelson, Cox: Lehninger, Principles in Biochemistry
- Takahashi, Schnarr: Investigation of RecA polynucleotide interactions from the measurement of LexA repressor cleavage kinetics



- Michel: After 30 Years of Study, the Bacterial SOS Response Still Surprises Us
- Fernandez De Henestrosa et. al.: Identification of additional genes belonging to the LexA regulon in Escherichia coli.