Pancreatic Cancer Research and HMGB1 Signaling Pathway

Haijun Gong*, Paolo Zuliani*, Anvesh Komuravelli*, James R. Faeder#, Edmund M. Clarke*

* Carnegie Mellon
# University of Pittsburgh
The Hallmarks of Cancer

D. Hanahan and R. A. Weinberg
Cell, Vol. 100, 57–70, January 7, 2000
1. Introduction
   • HMGB1 Protein
   • Important Signaling Pathways

2. Model Building
   • BioNetGen Model
   • Simulation Results

3. Model Checking
   • Statistical Model Checking
   • Verification of HMGB1 model

4. Conclusions
The Protein HMGB1

- High-Mobility Group Protein 1 (HMGB1):
  - DNA-binding protein and regulates gene transcription
  - released from damaged or stressed cells, etc.

- HMGB1 activates RAGE or TLR2/4
  - RAGE: Receptor for Advanced Glycation End products.
  - TLR: Toll-like receptor

- RAGE/TLR activation can activate NFκB and RAS signaling pathways which causes inflammation or tumorigenesis.
Experiments with pancreatic cancer cells:

- Overexpression of HMGB1/RAGE is associated with diminished apoptosis, and longer cancer cell survival time.
- Knockout of HMGB1/RAGE leads to increased apoptosis, and decreased cancer cell survival.
We use the BioNetGen language (http://bionetgen.org) to describe the crosstalk of important signaling pathways activated by HMGB1.

- We focus on the p53, RAS, NFkB & RB-E2F signaling pathways.
- How the expression level of HMGB1 influences the cell’s fate.

We use statistical model checking to formally verify behavioral properties expressed in temporal logic:

- Can express quantitative properties of systems
- Scalable, can deal with large models
P53-RAS-RB Crosstalk Model

- **First** computational model of HMGB1 signal transduction in tumorigenesis.
- Focus on the crosstalk of p53, RAS, & RB signaling pathways.
- More details in the paper “Analysis and Verification of the HMGB1 Signaling Pathway” published in *BMC Bioinformatics 11 (Suppl 7) (2010)*
P53 is a tumor suppressor, is mutated in more than 50% of cancers.

Functions of P53:
1. **Induces cell cycle arrest**: P21, etc.
2. **DNA repair**
3. **Initiates apoptosis** – Programmed Cell Death: Bax, etc.

- **Negative feedback loop**:
  \[ \text{PI3K} \rightarrow \text{PIP3} \rightarrow \text{AKT} \rightarrow \text{MDM2} \rightarrow \text{p53} \rightarrow \text{MDM2} \]

- **Positive feedback loop**:
  \[ \text{p53} \rightarrow \text{PTEN} \rightarrow \text{PIP3} \rightarrow \text{AKT} \rightarrow \text{MDM2} \rightarrow \text{p53} \]
1. Activation of RAS signaling causes cell growth and survival.
2. RAS family has **three members**: HRAS, KRAS, NRAS.
3. KRAS mutations are found in more than 90% of pancreatic cancers

- RAGE $\rightarrow$ RAS $\rightarrow$ RAF $\rightarrow$ MEK $\rightarrow$ ERK1/2 $\rightarrow$ TFs $\rightarrow$ Cyclin D $\rightarrow$ Cell-cycle progression
- RAS $\rightarrow$ PI3K $\rightarrow$ PIP3 $\rightarrow$ AKT $\rightarrow$ MDM2 $\rightarrow$ Apoptosis
RB-E2F pathway

- Regulates the G1-S phase transition in the cell cycle.

1. E2F is an oncprotein, activates the transcription of Cyclin E, and it is modulated by RB.

2. RB is a tumor suppressor: prevents the replication of damaged DNA.

3. Cyclin D-CDK4 phosphorylates RB, leading to the activation of E2F.

- CyclinD → RB → E2F → CyclinE → S Phase
31 molecular species
59 reactions

Blue: tumor suppressor
Red: oncoprotein/gene
P53-NFkB-RAS-RB Crosstalk Model

- Crosstalk of p53, NFkB, RAS, & RB signaling pathways.
- NFkB protein is involved in inflammation, cell proliferation and apoptosis.
- NFkB is a transcription factor for the pro-apoptotic gene p53, for anti-apoptotic genes Bcl-XL and for the cell-cycle regulatory proteins Myc and Cyclin D.
In resting cells \( \text{IkB} \) exists only in the cytoplasm, bound to \( \text{NFkB} \).

\( \text{IKK} \) (\( \text{IkB} \) kinase) can phosphorylate \( \text{IkB} \) to release \( \text{NFkB} \).

Free \( \text{NFkB} \) enter the nucleus to activate the expression of \( \text{A20} \), \( \text{IkB} \), \( \text{P53} \), \( \text{Cyclin D} \), \( \text{Myc} \).

Overexpression of \( \text{NFkB} \) is common in the pancreatic cancer.

Inflammation

Apoptosis

NFkB pathway

2 negative feedback loops:

1. \( \text{TLR} \rightarrow \text{IKK} \rightarrow \text{IkB} \rightarrow \text{NFkB} \)
2. \( \text{NFkB} \rightarrow \text{A20} \rightarrow \text{IKK} \rightarrow \text{IkB} \rightarrow \text{NFkB} \)
begin molecule types

A(b,Y~U~P) # A has a component Y which
    # can be labeled as U (unphosphorylated)
    # or P (phosphorylated)

B(a)

end molecule types

begin reaction rules

A(b) + B(a) <-> A(b!1).B(a!1)

A(Y~U) -> A(Y~P)

end reaction rules

Ordinary Differential Equations and Stochastic simulation

BioNetGen

- **Two Events:** PIP3 phosphorylates AKT, and AKT dephosphorylates.

**begin species**

<table>
<thead>
<tr>
<th>Species</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKT (d~U)</td>
<td>1e5</td>
</tr>
<tr>
<td>AKT (d~p)</td>
<td>0</td>
</tr>
</tbody>
</table>

**begin parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>1.2e-7</td>
</tr>
<tr>
<td>d</td>
<td>1.2e-2</td>
</tr>
</tbody>
</table>

**end species**

**begin reaction_rules**

1. $\text{PIP}(c~p) + \text{AKT}(d~U) \rightarrow \text{PIP}(c~p) + \text{AKT}(d~p)$ \quad k
2. $\text{AKT}(d~p) \rightarrow \text{AKT}(d~U)$ \quad d

**end reaction_rules**

- The corresponding **ODE** is:

$$[\text{AKT}(d~p)](t)' = k \cdot [\text{PIP}(c~p)](t) \cdot [\text{AKT}(d~U)](t) - d \cdot [\text{AKT}(d~p)](t)$$
Simulations (I)

- Baseline simulation of p53, MDM2, Cyclin D/E in response to HMGB1 release: ODE vs stochastic simulation
Simulations (II)

- Baseline simulation of NFκB, IkB, IKK and A20 in response to HMGB1 release.
Simulations (III)

- Overexpression of HMGB1 leads to increase of E2F and Cyclin D/E, decrease of p53.

- Overexpression of AKT represses p53 level.
Simulations (IV)
Simulations (V)

- IKK overexpress in many cancer cells, it promotes NFkB’s transcription activity and accelerate cell proliferation.
- Overexpression of NFkB is common in pancreatic cancer.

![Graphs showing IKK overexpression and NFkB activity](image)
Bounded Linear Temporal Logic

- **Bounded Linear Temporal Logic (BLTL):** Extension of LTL with time bounds on temporal operators.
- \( F^t a \) – “a will be true in the Future *within time* \( t \)”
- \( G^t a \) – “a will be Globally true *between time 0 and t*”
- For example: “# of AKTp reach 4,000 within 20 minutes?” – \( F^{20} (AKTp \geq 4,000) \)
- Let \( \sigma = (s_0, t_0), (s_1, t_1), \ldots \) be an execution of the model
  - along states \( s_0, s_1, \ldots \)
  - the system stays in state \( s_i \) *for time* \( t_i \)
- \( \sigma^i \): Execution trace starting at state \( i \).
Verification of BioNetGen Models

- Given a stochastic BioNetGen model $\mathcal{M}$, Temporal property $\Phi$, and a fixed $0<\theta<1$, we ask whether $P_{\geq \theta}(\Phi)$ or $P_{<\theta}(\Phi)$.

- For example: “could AKTp reach 4,000 within 20 minutes, with probability at least 0.99?” : $P_{\geq 0.99}(F^{20}(AKTp \geq 4,000))$

- Does $\mathcal{M}$ satisfy $\phi$ with probability at least $\theta$? $\mathcal{M} \models P_{\geq \theta}(\phi)$

- Draw a sample of system simulations and use Statistical Hypothesis Testing: Null vs. Alternative hypothesis
  $H_0 : \mathcal{M} \models P_{\geq \theta}(\phi)$  $H_1 : \mathcal{M} \models P_{<\theta}(\phi)$
Statistical Model Checking of biochemical models: $\mathcal{M} \models P_{\geq \theta}(\Phi)$?
Bayes Factor

- \( X = (x_1, \ldots, x_n) \) a sample of Bernoulli random variables
- Prior probabilities \( P(H_0), P(H_1) \) strictly positive, sum to 1
- Ratio of Posterior Probabilities:
  \[
  \frac{P(H_0|X)}{P(H_1|X)} = \frac{P(X|H_0) \cdot P(H_0)}{P(X|H_1) \cdot P(H_1)}
  \]

Bayes Factor B

- Fix threshold \( T \geq 1 \) and prior probabilities \( P(H_0), P(H_1) \).
  Continue sampling until
  - Bayes Factor B > \( T \): Accept \( H_0 \)
  - Bayes Factor B < \( 1/T \): Reject \( H_0 \)
SMC Algorithm

Require: Property $P_{\geq \theta}(\Phi)$, Threshold $T \geq 1$, Prior density $g$

$n := 0$ \{number of traces drawn so far\}

$x := 0$ \{number of traces satisfying $\Phi$ so far\}

repeat
  \[
  \sigma := \text{draw a sample trace from BioNetGen (iid)}
  \]
  \[
  n := n + 1
  \]
  if $\sigma \models \Phi$ then
    \[
    x := x + 1
    \]
  endif
  \[
  B := \text{BayesFactor}(n, x, g)
  \]
until $(B > T \lor B < 1/T)$

if $(B > T)$ then
  return “$H_0$ accepted”
else
  return “$H_0$ rejected”
endif
Verification (I)

- Overexpression of HMGB1 will induce the expression of cell regulatory protein CyclinE.
- We model checked the formula with different initial values of HMGB1, the probability error is 0.001.

\[ P_{\geq 0.9} F_{600}^{} (\text{CyclinE} > 900) \]

<table>
<thead>
<tr>
<th>HMGB1</th>
<th># samples</th>
<th># Success</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^2$</td>
<td>9</td>
<td>0</td>
<td>False</td>
</tr>
<tr>
<td>$10^3$</td>
<td>55</td>
<td>16</td>
<td>False</td>
</tr>
<tr>
<td>$10^6$</td>
<td>22</td>
<td>22</td>
<td>True</td>
</tr>
</tbody>
</table>
Verification (II)

- *P53 is expressed at low levels in normal human cells.*
- \( P_{\geq 0.9} F^t \left( G^{900} \left( p53 < 3.3 \times 10^4 \right) \right) \)

<table>
<thead>
<tr>
<th>t(min)</th>
<th># Samples</th>
<th># Success</th>
<th>Result</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>53</td>
<td>49</td>
<td>True</td>
<td>597.59</td>
</tr>
<tr>
<td>500</td>
<td>23</td>
<td>22</td>
<td>True</td>
<td>271.76</td>
</tr>
<tr>
<td>600</td>
<td>22</td>
<td>22</td>
<td>True</td>
<td>263.79</td>
</tr>
</tbody>
</table>
Verification (III)

- Expression level of HMGB1 influence the 1st peak of p53’s level.

\[ P_{\geq 0.9} \mathbf{F}^{100}(p53 \geq a \& \mathbf{F}^{100}(p53 \leq 4 \times 10^4)) \]

<table>
<thead>
<tr>
<th>HMGB1</th>
<th>a ( x 10^4 )</th>
<th># Samples</th>
<th># Success</th>
<th>Result</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^3</td>
<td>5.5</td>
<td>20</td>
<td>3</td>
<td>False</td>
<td>29.02</td>
</tr>
<tr>
<td>10^2</td>
<td>5.5</td>
<td>22</td>
<td>22</td>
<td>True</td>
<td>19.65</td>
</tr>
<tr>
<td>10^2</td>
<td>6.0</td>
<td>45</td>
<td>12</td>
<td>False</td>
<td>56.27</td>
</tr>
<tr>
<td>10</td>
<td>6.0</td>
<td>38</td>
<td>37</td>
<td>True</td>
<td>41.50</td>
</tr>
</tbody>
</table>
Verification (IV)

- HMGB1 can activate PI3K, RAS and AKT in large quantities.
- Let $\text{PI3Kr}$, $\text{RASr}$, and $\text{IKKr}$ be the fraction of activated molecules of PI3K, RAS, and IKK, respectively.
- We model checked the formula:

$$P_{\geq 0.9} F_t G^{180} (\text{PI3Kr} > 0.9 \land \text{RASr} > 0.8 \land \text{IKKr} > 0.6)$$

<table>
<thead>
<tr>
<th>$t$ (min)</th>
<th># Samples</th>
<th># Success</th>
<th>Result</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>9</td>
<td>0</td>
<td>False</td>
<td>21.27</td>
</tr>
<tr>
<td>110</td>
<td>38</td>
<td>37</td>
<td>True</td>
<td>362.19</td>
</tr>
<tr>
<td>120</td>
<td>22</td>
<td>22</td>
<td>True</td>
<td>214.38</td>
</tr>
</tbody>
</table>
Coding oscillations of NFkB in temporal logic

Let \( R \) be the fraction of NFkB molecules in the nucleus

\[ P_{\geq 0.9} F_t (R \geq 0.65 \& F_t (R < 0.2 \& F_t (R \geq 0.2 \& F_t (R < 0.2)))) \]

<table>
<thead>
<tr>
<th>HMGB1</th>
<th>t (min)</th>
<th># Samples</th>
<th># Success</th>
<th>Result</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(10^2)</td>
<td>45</td>
<td>13</td>
<td>1</td>
<td>False</td>
<td>76.77</td>
</tr>
<tr>
<td>(10^2)</td>
<td>60</td>
<td>22</td>
<td>22</td>
<td>True</td>
<td>111.76</td>
</tr>
<tr>
<td>(10^2)</td>
<td>75</td>
<td>104</td>
<td>98</td>
<td>True</td>
<td>728.65</td>
</tr>
<tr>
<td>(10^5)</td>
<td>30</td>
<td>4</td>
<td>0</td>
<td>False</td>
<td>5.76</td>
</tr>
</tbody>
</table>
Conclusions

- Computational model qualitatively confirmed the previous HMGB1 experimental phenomena.
- Our simulations predict a dose-dependent p53, CyclinE, and NFkB response curve to increasing HMGB1 stimulus.
- Statistical Model Checking automatically validate our model with respect to known experimental results.
Future Work

- Parameter estimation
- Combine Machine Learning (Bayesian Network) and Model Checking to infer Gene Regulatory Network
- Multi-cellular systems
- Pancreatic stellate cells
Acknowledgments

- This work supported by the NSF Expeditions in Computing program

- Thanks to Michael T. Lotze (University of Pittsburgh) for calling our attention to HMGB1

- Thanks to Marco E. Bianchi (Università San Raffaele) for discussions on HMGB1
Thank you!

Questions?