Some very hard problems in nature (biology–biochemistry) “solved” using physical algorithms that reduce the hardness

“Problems”

Search optimization
Hill climbing—energy reduction
Allocation of resources
Self assembly
Reversible computation
Satisfiability
Controllers for nanomachines

“Algorithms”

Cooperativity
Heterogeneity
Stochasticity

add your favorite problem

PENN HUNT PROJECT
September 18, 2008

Harvey Rubin MD, PhD
University of Pennsylvania
**Title:** Some Very Hard Problems in Nature (Biology-biochemistry) 'Solved' Using Physical Algorithms that Reduce the Hardness

**Performing Organization:** University of Pennsylvania, Department of Electrical and Systems Engineering, Philadelphia, PA, 19104

**Approved for public release; distribution unlimited**
Cooperativity at the monomolecular level binding of B or C to the common partner A affects binding of the other.

**Figure 1.** Thermodynamic cycles and cooperativity. (a) Hypothetical set of bimolecular complexes between component A and two other components (B and C), with the rate constants, equilibrium constants and free energies for complex formation. (b) A thermodynamic cycle for formation of the ternary complex ABC by two different possible routes: either B binds first, or C binds first. There are four equilibrium constants that describe the formation of the various complexes. Because they converge on the common product ABC, the thermodynamics must be independent of the pathway chosen around the cycle, and constraints are placed on the relative values of the equilibrium constants and hence the free energies. The thermodynamic coupling free energy ($\Delta G^{coupling}$) gives the difference between binding of one component in the presence of the other. (c) Definition of cooperativity in terms of binding of B in the presence or absence of C. The two vertical binding reactions are gray to emphasize the comparison of $\Delta G^{1}_{A}$ and $\Delta G^{2}_{A}$. If B binds better in the presence of C, the binding is cooperative. If B binds worse in the presence of C, the binding is anticooperative. In the third case, binding of B is independent of C, and there is no cooperativity.
1. Complex interactions among identical ligands binding to multiple sites on an oligomeric protein—oxygen binding to hemoglobin.

Homotropic allosteric regulators—e.g. O2

Heterotropic allosteric regulators—e.g. 2,3 BPG

2. The thermodynamics of macromolecular conformational transitions—protein folding or nucleic acid helix-coil transitions.

3. The thermodynamics of forming multicomponent complexes—multimeric complexes, surface interactions, cellular communication, organism organization, multicellular dynamics, social structures

Cooperativity and biological complexity
Adrian Whitty nature chemical biology volume 4 number 8 august 2008
Interaction of Hemoglobin with Three Ligands: Organic Phosphates and the Bohr Effect

(Haldane coefficient/2,3-diphosphoglycerate/linkage equations/allosteric effect/oxygen binding)

RUTH E. BENESCH AND HARVEY RUBIN
Department of Biochemistry, Columbia University, College of Physicians & Surgeons, New York, N.Y.

Communicated by Harden M. McConnell, April 15, 1975

ABSTRACT. The assumption that the Bohr coefficient \( \frac{\Delta \log p_O}{\Delta pH} \) is equal to the Haldane coefficient \( \Delta E^{\Pi} \) of hemoglobin is shown to be incorrect in the presence of allosteric effectors such as 2,3-diphosphoglycerate. The theoretical relation between the two coefficients in the presence of 2,3-diphosphoglycerate is derived. Experimental data on the variation of both coefficients with diphosphoglycerate concentration are presented and shown to be in agreement with prediction.

Therefore, the liberation of diphosphoglycerate on oxygenation must lead to an increase in its activity with increasing oxygenation, and Eqs. 1 and 2 cannot any longer apply.

It is the purpose of this paper to examine the relation between the Bohr and Haldane effects in the presence of 2,3-diphosphoglycerate.

RESULTS AND DISCUSSION

Box 1 A timeline showing evolution of allostery as a concept

The timeline (Fig. 1) includes some of the key experiments and realizations of the field and some insight into how the mentality of the field has shifted.

1903—The Bohr effect (sigmoidal binding curve of hemoglobin to \( O_2 \) was observed).
1910—A. Hill formulates the Hill equation to describe the sigmoidal binding of \( O_2 \) to hemoglobin.
1958—First X-ray structure (sperm whale myoglobin) solved by M. Perutz and Sir J. Cowdery Kendrew.79
1950s—Repulsion of gene expression, covalent modification of enzyme activity, and feedback inhibition of enzymes are discovered79.
1963—J. Monod renames regulatory sites 'allosteric sites.'
1966—D. Koshland, G. Nemethy and R. Filmer propose the sequential model for allosteric transitions (KNF model)80.
1984—Allosteric regulation in the absence of conformational change is proposed12.
1980s—Protein folding studies lead to the concept that proteins exist in different conformations in an "energy landscape"81.
1990s—Mutations, covalent modifications and changes in conditions such as pH are included as allosteric effectors.
1999—Allosteric networks in the PDZ domain proposed by R. Ranganathan50.
2006—Negative allostery reported in the absence of conformational changes54.

Figure 1. Timeline showing the discovery and progression of the concept of allosteric regulation in proteins.
Logarithmic scale of $k_{\text{cat}}$ and $k_{\text{non}}$ values for representative reactions at 25 °C. The length of each vertical bar represents the rate enhancement by each enzyme.

The Depth of Chemical Time and the Power of Enzymes as Catalysts

R. WOLFENDEN AND M.J. SNIDER

Acc. Chem. Res. 2001, 34, 938-945
How does “Biology” cope?

“After total war can come total living”

Mutually Assured Destruction: Cold War exhibit at the Smithsonian
Stringent response and growth control

- Triggered by adverse conditions, e.g. starvation

- Transcription control (p)ppGpp:
  - Lack of nutrients
  - Stalled ribosomes
  - ppGpp synthesis
  - Reprogramming of transcription

- Translation shutdown:
  - Proteases
  - (p)ppGpp involved
  - Activation of toxin-antitoxin modules
  - Toxin reversibly disables ribosomes
Stringent response and growth control

- Triggered by adverse conditions, e.g. starvation

  - Transcription control (p)ppGpp:
    - Lack of nutrients
    - Stalled ribosomes
    - ppGpp synthesis
    - Reprogramming of transcription

  - Translation shutdown
    - Proteases
    - (p)ppGpp involved
    - Activation of toxin-antitoxin modules
    - Toxin reversibly disables ribosomes
Stringent response and growth control

- Triggered by adverse conditions, e.g. starvation

  - Transcription control \((p)ppGpp\):
    - Lack of nutrients
    - Stalled ribosomes
    - \(ppGpp\) synthesis
    - Reprogramming of transcription

  - Translation shutdown
    - Proteases
    - \((p)ppGpp\) involved
    - Activation of toxin-antitoxin modules
    - Toxin reversibly disables ribosomes
Stringent response and growth control

- Triggered by adverse conditions, e.g. starvation

Transcription control (p)ppGpp:
- Lack of nutrients
- Stalled ribosomes
- ppGpp synthesis
- Reprogramming of transcription

Translation shutdown:
- Proteases
- (p)ppGpp involved
- Activation of toxin-antitoxin modules
- Toxin reversibly disables ribosomes
Stringent response and growth control

• Triggered by adverse conditions, e.g. starvation

• Transcription control (p)ppGpp:
  - Lack of nutrients
  - Stalled ribosomes
  - ppGpp synthesis
  - Reprogramming of transcription

• Translation shutdown
  - Proteases
  - (p)ppGpp involved
  - Activation of toxin-antitoxin modules
  - Toxin reversibly disables ribosomes

![Diagram showing the STRINGENT response and growth control

- TRANSCRIPTION
- TRANSLATION
- GROWTH
- ppGpp
- Lon
- Toxins
- RAC
- NUTRIENT AVAILABILITY]
The Stringent Response is mediated by two opposing $\text{Rel}_{\text{Mtb}}$ activities which must be tightly regulated.

1) pppGpp synthesis:

\[
\begin{align*}
p-p-p-G + p-p-p-A & \leftrightarrow p-p-p-G-p-p + p-A \\
GTP + ATP & \leftrightarrow G5 + AMP
\end{align*}
\]

2) pppGpp hydrolysis:

\[
\begin{align*}
p-p-p-G-p-p & \leftrightarrow \text{PPi} + p-p-p-p-G
\end{align*}
\]

pppGpp alters RNAP kinetics and mediates the transcriptional response to environmental conditions to which Mtb is exposed.
The RAC Allosterically Activates Transferase Activity

\[ \text{RAC} = \text{Rel}_{Mtb} \text{ Activating Complex} \]

Ribosome•Uncharged tRNA•mRNA

Rel\textsubscript{Mt}b (Basal Level)

<table>
<thead>
<tr>
<th>( K_{\text{ATP}} ) (mM)</th>
<th>( K_{\text{GTP}} ) (mM)</th>
<th>( k_{\text{cat}} ) (s(^{-1}))</th>
<th>( k_{\text{cat}}/K_{\text{ATP}} ) (mM(^{-1}) s(^{-1}))</th>
<th>( k_{\text{cat}}/K_{\text{GTP}} ) (mM(^{-1}) s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>1.4</td>
<td>1.2</td>
<td>0.6</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Rel\textsubscript{Mt}b + Ribosome•UtRNA•mRNA

<table>
<thead>
<tr>
<th>( K_{\text{ATP}} ) (mM)</th>
<th>( K_{\text{GTP}} ) (mM)</th>
<th>( k_{\text{cat}} ) (s(^{-1}))</th>
<th>( k_{\text{cat}}/K_{\text{ATP}} ) (mM(^{-1}) s(^{-1}))</th>
<th>( k_{\text{cat}}/K_{\text{GTP}} ) (mM(^{-1}) s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>24.7</td>
<td>54.8</td>
<td>79.6</td>
</tr>
</tbody>
</table>

\[ \text{5’..AUGCCGACGUACAGUUUGUUCGGGC...3’} \]
Heterogeneity even within a single molecule

Figure 1: Summary of Rel\textsubscript{Mtb} truncated. Full-length Rel\textsubscript{Mtb} protein is at the top followed by the different truncated proteins. Amino acid numbers are at the beginning and end of each fragment and corresponding activity is listed below. 87-187 overlapping site is noted in the full-length Rel\textsubscript{Mtb}.
Cooperativity, heterogeneity, stochasticity
Another example: Controllers for nanomachines
Aerobic and anaerobic respiratory chain in Mtb

- Fumarate Reductase
  - FrdABCD (Rv1552-Rv1555)

- Succinate:Menaaquinone Oxidoreductase
  - sdhABCD (Rv3316-Rv3319)

- Electron-transferring Flavoproteins

- NADH:Menaaquinone Oxidoreductases
  - Type I: nuoABCDEFGHJKLMN (Rv3145-Rv3158)
  - Type II: ndh (Rv1854c)
  - Type II: ndhA (Rv0392c)

- Menaquinone Pool
  - menABCDEG, ubiE (Rv0534c, Rv0548c, Rv0553, Rv0555, Rv0542c, Rv3853, Rv0558)

- Nitrate Reductase
  - Fused: narX (Rv1736c)
  - Multisubunit: narGHIJ (Rv1161-Rv1164)

- Cytochrome bd Oxidase
  - cydABCD (Rv1620c-1623c)

- Cytochrome c Oxidase (aa₃)
  - ctaBCDE (Rv1451, Rv2200c, Rv3043c, Rv2193)

- ATP synthase

Aerobic Pathway

Anaerobic Pathway
Electrons enter the chain through NADH oxidoreductase

Plot of the NADH-Q2 reductase reaction with varying Q concentrations and fixed concentrations of NADH. Lineweaver-Burk plot (inset), slopes (Vmax/Km) of the lines are not affected by NADH concentration—ping pong mechanism

\[ K_{m}^{NADH} = 42 \text{ uM}, \quad K_{m}^{Q2} = 12.5 \text{ uM}, \quad V_{\text{max}} = 26 \text{ unit mg}^{-1} \]
• NDH-2 catalyzes the following two electron transfer reactions:

\[ \text{Ndh(Fl}_{\text{ox}}) + \text{NADH} \rightarrow \text{Ndh(Fl}_{\text{red}}) + \text{NAD}^+ \] \hspace{1cm} (eq1)

\[ \text{Ndh(Fl}_{\text{red}}) + \text{Q} \rightarrow \text{Ndh(Fl}_{\text{ox}}) + \text{QH}_2 \] \hspace{1cm} (eq2)

\[ E \rightarrow (EA \leftrightarrow FP) \rightarrow F \rightarrow (FB \leftrightarrow FR) \rightarrow E \]

\[ k_1 \quad k_{-1} \quad k_2 \quad k_{-2} \quad k_3 \quad k_{-3} \quad k_4 \quad k_{-4} \]

\( E \) is MtB NDH-2, \( A \) is NADH and \( B \) is the quinone
Phenothiazine inhibition of Mtb respiration.

(A) TPZ inhibition of NADH-dependent oxygen consumption by Mtb membranes measured with a Clark-type oxygen electrode. Respiration was initiated by the addition of 10 mM NADH and arrested upon the addition of 1mMTPZ. Addition of 10mM ascorbate and 1 mM TMPD produced an immediate resumption of respiration.
A 3D model of *E. coli* Ndh according to Schmid and Gerloff (2004). Putative flavin-, NADH-, and membrane-binding domains are shown in ovals.
A drug for dormant TB

<table>
<thead>
<tr>
<th>Drug</th>
<th>MBC(mg/L)</th>
<th>Log-phase</th>
<th>6-week-starved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rifampin</td>
<td>&lt;0.625</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Trifluoperazine</td>
<td>10~20</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>10~20</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Isoniazid</td>
<td>&lt;0.625</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Ethionamide</td>
<td>&lt;0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Capreomycin sulfate</td>
<td>0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Amikacin sulfate</td>
<td>&lt;0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Thiacetazone</td>
<td>&lt;0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Ethambutol</td>
<td>0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Streptomycin sulfate</td>
<td>&lt;0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>p-aminosalicylic acid</td>
<td>&lt;0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Ofloxacin</td>
<td>&lt;0.625</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Tetracycline</td>
<td>10~20</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Cycloserine</td>
<td>10~20</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Erythromycin</td>
<td>40</td>
<td>&gt;160</td>
<td></td>
</tr>
<tr>
<td>Dapsone</td>
<td>&gt;40</td>
<td>&gt;160</td>
<td></td>
</tr>
</tbody>
</table>

MBC$_{99}$s of 17 Drugs for Log-phase and 6-week-starved M.
tuberculosisH37Rv by cfu counts.
We shall destroy the respiratory chain by recognizing the heterogeneity of its parts!

We shall go on to the end, we shall fight in France, we shall fight on the seas and oceans, we shall fight with growing confidence and growing strength in the air, we shall defend our Island, whatever the cost may be, we shall fight on the beaches, we shall fight on the landing grounds, we shall fight in the fields and in the streets, we shall fight in the hills; we shall never surrender.
WSC June 4, 1940
Can molecular computing say anything based on irreversible nature of computation

The Fundamental Physical Limits of Computation

*What constraints govern the physical process of computing? Is a minimum amount of energy required, for example, per logic step? There seems to be no minimum., but some other questions are open.*

by Charles H. Bennett and Rolf Landauer

A Fredkin Gate: Logically reversible with no energy limit on the computation

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>-&gt;</th>
<th>A'</th>
<th>B'</th>
<th>C'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

CAB is a piece of DNA that we can synthesize
A NAND gate

**Figure 2**
Why reversible?

Minimal energy expense

Detection and correction of intrusion

Error checking by reversing computation to recreate inputs

Bidirectional debugging
In principle it can take minimal energy to go through a biochemical gate

\[ \text{DNA}_n + \text{dNTP} \rightleftharpoons \text{DNA}_{n+1} + \text{PPi} \]

\[ \Delta G = kt \ln[\text{dNTP}/\text{PPi}] \]

If dNTPs are just 1% over the equilibrium value:

\[ \Delta G = kt \ln[10.1/10] \quad \text{or about } 0.01kT \]

a modification of an idea in Bennett and Landaur's Sci. Am paper—suggested using RNA
We synthesized the oligonucleotides and ran the reactions.

The gate works in the lab

\[ \begin{align*}
0 \times 1 \text{ (lane2)} &\rightarrow \text{ no product (lane5)} \\
1 \times 0 \times 0 \times 1 \text{ (lane4)} &\rightarrow \text{ no product (lane5)} \\
1 \times 0 \times 0 \times (\text{lane7}) &\rightarrow 0 \times 1 \text{ (lane9)} \\
0 \times 1 \text{ (lane2)} &\rightarrow \text{ no product (lane10)}
\end{align*} \]
How fast could one go through one gate?

$\tau_{1/2}$ annealing: 3 sec.

DNA polymerization rate: 15 bases/sec

For 60 bases pair input: 10 sec
Some very hard problems in nature (biology–biochemistry) “solved” using physical algorithms that reduce the hardness

<table>
<thead>
<tr>
<th>“Problems”</th>
<th>“Algorithms”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search optimization</td>
<td>Cooperativity</td>
</tr>
<tr>
<td>Hill climbing—energy reduction</td>
<td>Heterogeneity</td>
</tr>
<tr>
<td>Allocation of resources</td>
<td>Stochasticity</td>
</tr>
<tr>
<td>Self assembly</td>
<td></td>
</tr>
<tr>
<td>Reversible computation</td>
<td></td>
</tr>
<tr>
<td>Satisfiability</td>
<td></td>
</tr>
<tr>
<td>Controllers for nanomachines</td>
<td></td>
</tr>
</tbody>
</table>

add your favorite problem
Acknowledgements

School of Medicine
Mei Wang
Edward Weinstein
Andrew Avarbock
Michael Buckstein
Jamaine Davis
Marcin Imielinski
Jeh Shin The
Norman Schechter
Takahiro Yano

Penn Computer Science
Vijay Kumar
Adam Halasz
Oleg Sokolsky
Calin Belta
George Pappas

Funding
NIH
DARPA
NSF
Global Alliance for TB Drug Development