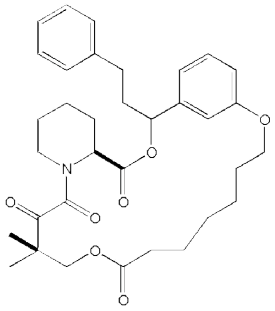


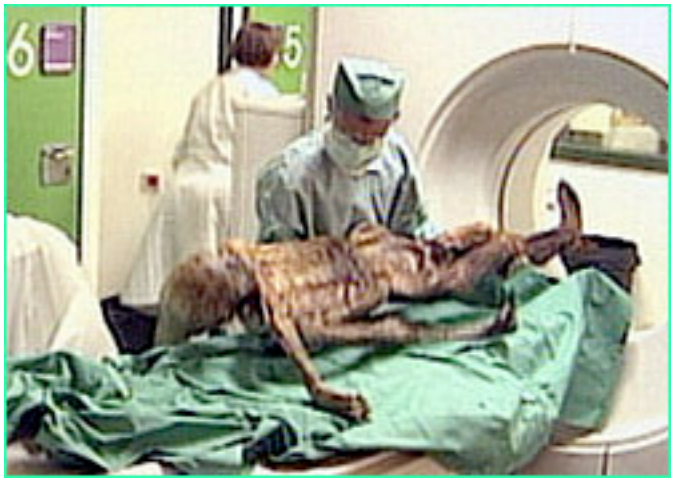
Computational Drug Discovery

Guha. January 10, 2006

Two Revolutions

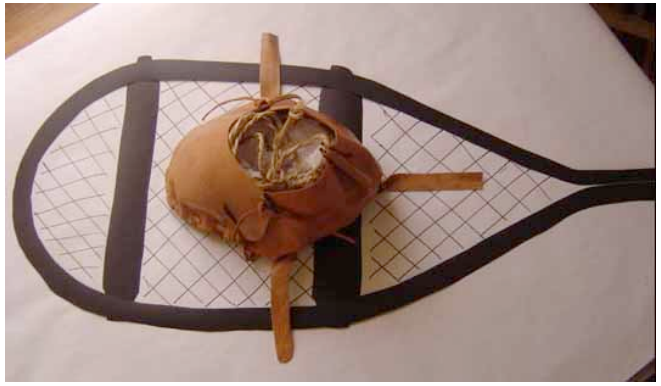
Guha. January 10, 2006

A Corpse in the Alps



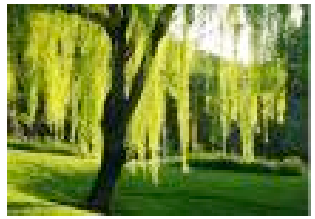
Why interesting?

His Possessions



Search for Drugs Not New

- Traditional Chinese medicine and Ayurveda both several thousand years old
  - Many compounds now being studied
- Aspirin's chemical forefather known to Hippocrates



- Even inoculation at least 2000 years old

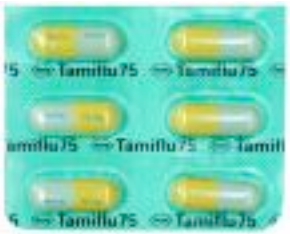
- And, of course, many useless drugs too

More Concerted Efforts

- In 1796, Jenner finds first vaccine: cowpox prevents smallpox

- 1 century later, Pasteur makes vaccines against anthrax and rabies
- Sulfonamides developed for antibacterial purposes in 1930s
- Penicillin: the “miracle drug”
- 2nd half of 20th century: use of modern chemical techniques to create explosion of medicines

Towards Health



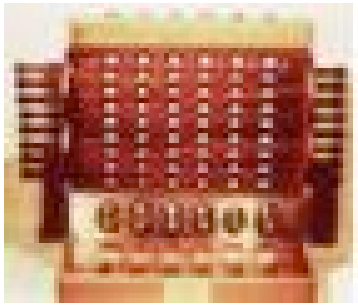
Not Enough

- AIDS and many cancers without cures despite billions of dollars spent
- Chronic ailments like blood pressure, arthritis, diabetes, etc. still need better therapies
- New problems like Mad Cow, SARS, and Avian flu emerging
- And old problems like infectious disease coming back, with antibiotic resistance growing
- At the same time, new lead molecules appearing less and less…

Computation's Progress

Fingers
(prehistoric)

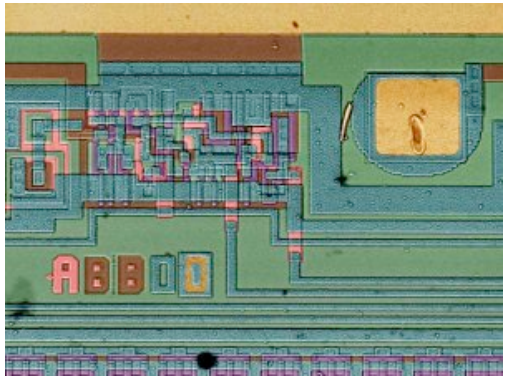
Abacus
(thousands of
years)



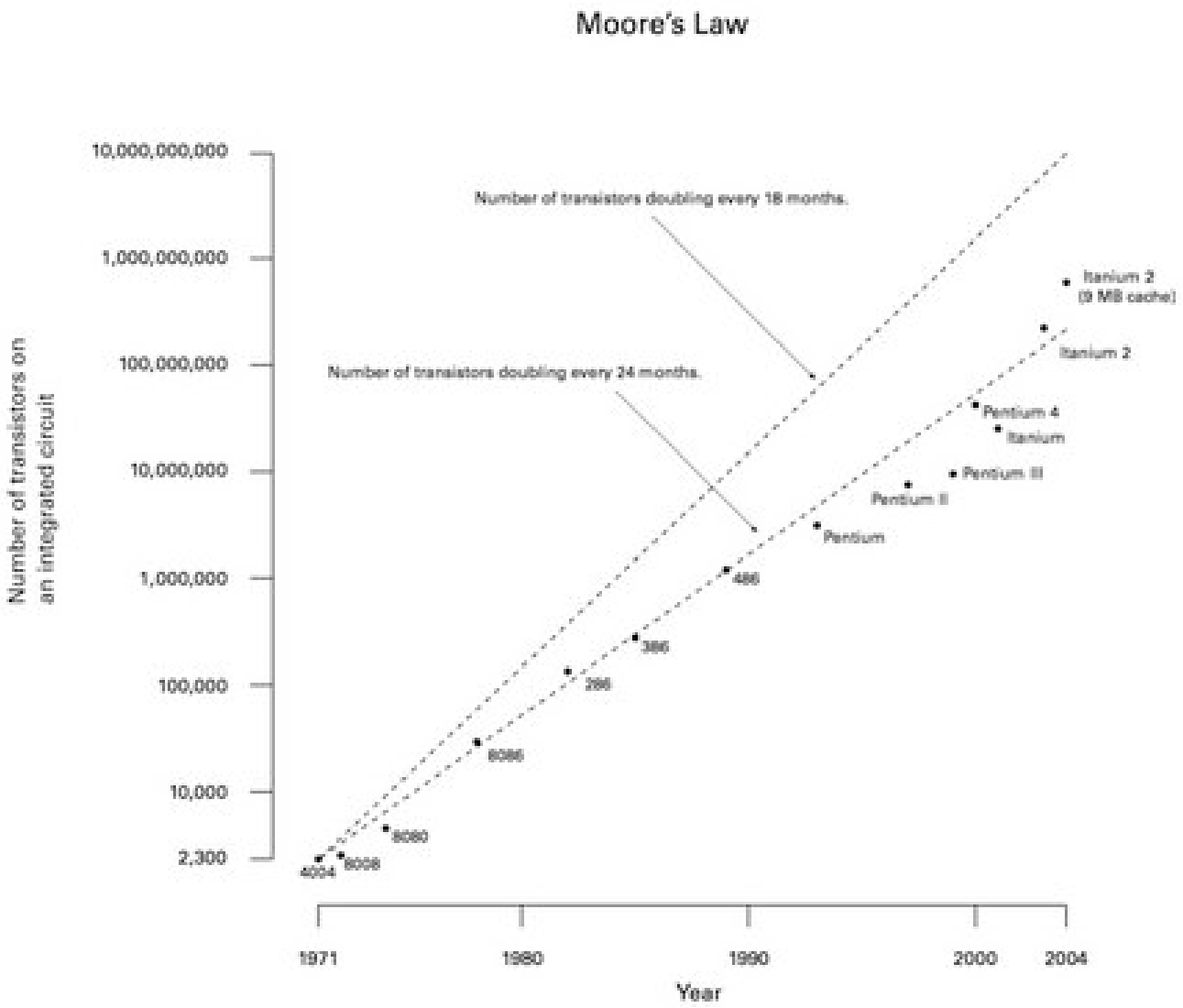
Mechanical
calculator (1623)

Even in beginning of 20th century, "computer" more a job title than a machine

Explosion of Progress



Moore's Law


Moore's Law
Number of transistors doubling every 18 months.
Number of transistors doubling every 24 months.
Number of transistors on an integrated circuit
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
2,300
4004
8008
8080
8086
286
386
486
Pentium
Pentium II
Pentium III
Pentium 4
Itanium
Itanium 2
Itanium 2 (9 MB cache)
1971
1980
1990
2000
2004
Year

Convergence

- Two great technological revolutions in last century
- In recent years, starting to come together
  - We will ignore computational tools that are only in support roles, like visualization
- Some computational methods for discovery now well established (like QSAR), others (more revolutionary) not yet integral part of mainstream discovery process

How Drugs Work (Briefly)

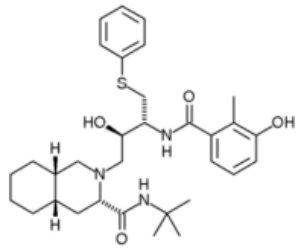
Guha. January 10, 2006

Small Molecule Drugs

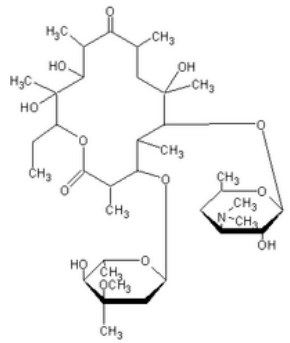
- Bind to a target
  - Can either be to a protein in one of our own cells, or can be to a foreign invader
- Cause some effect
  - Antagonists decrease activity
  - Agonists increase it

Examples

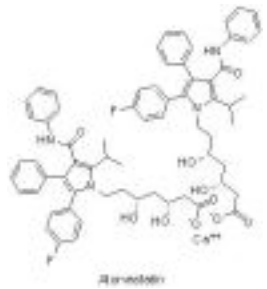
- Nelfinavir
  - Protease inhibitor used in treatment of HIV
  - Binds to HIV-1 and HIV-2 proteases, inhibiting them from cleaving viral protein



- Erythromycin
  - Antibiotic
  - Binds to bacterial ribosomes, stopping translation



- Statins
  - Class used to lower cholesterol
  - Inhibit HMG-CoA reductase, key enzyme in endogenous cholesterol production



The Goal

- First step is to find molecules that bind to target—it's hard
- That's not enough. Other requirements: should properly act as agonist and antagonist, should be something that can be synthesized, should be biomedically applicable (ADMET criteria)
- Each of those jobs is a challenge in and of itself

Why Compute

Guha. January 10, 2006

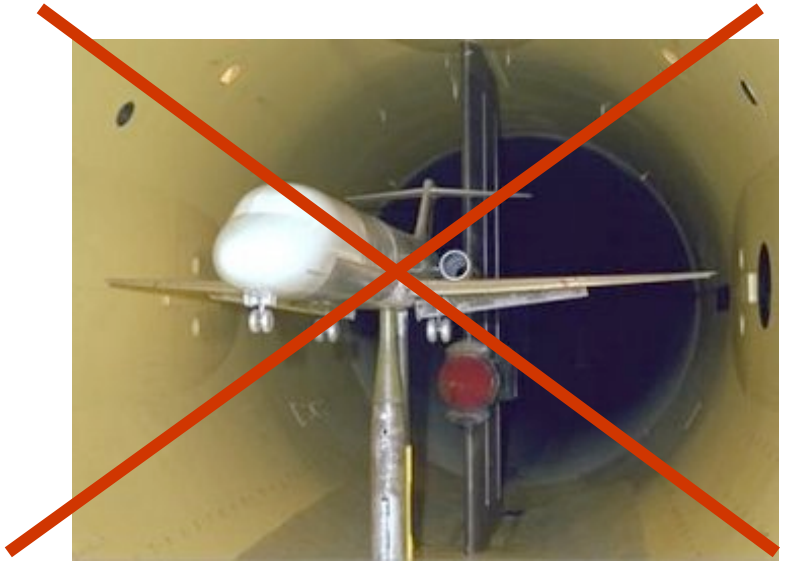
Status Quo Not OK

- Where's the cure for Alzheimer's? For the cold?
- Presently available small molecules target only ~500 of estimated 1 million human proteins
- Rate of new drugs going down: less approvals, more late stage failures
- Development of a new small molecule takes about 10 years and $1,000,000,000
- Unclear where next blockbuster drugs will come from

But Why Compute?

- To make possible the otherwise impossible
  - Can we design a molecule de novo and do initial toxicity tests without experiment?
  - Can we find new leads with just some time on a computer cluster instead of millions of dollars and years?
- Where does its potential come from?
  - Continue historical trend towards rationality, away from trial-and-error

Airplane Design



What's So Hard?

- Models
  - Molecular scale can't use simple macroscopic models
  - Need accuracy
  - But quantum mechanics too slow
- Processing power was lacking

Always Need Experiment

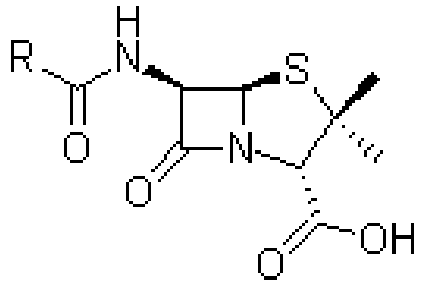
- Computation will not completely supplant experiment
  - Need data to test computational models
  - Humans are complex—can't simulate full effect of drug!
- Computation will reduce the amount of experiment by focusing it on the likeliest leads
  - Reduce time
  - Reduce cost
  - Increase results

Computational Methods in Context

Guha. January 10, 2006

1. Observation, Real World Discovery

- Classic example: penicillin discovered from mold experiments



- Go out, dig in the mud, collect samples, see if something works
  - FK506 an example
  - But we're not lucky enough

Mt. Tsukuba, where the mud that yielded FK506 was collected

2. Screening

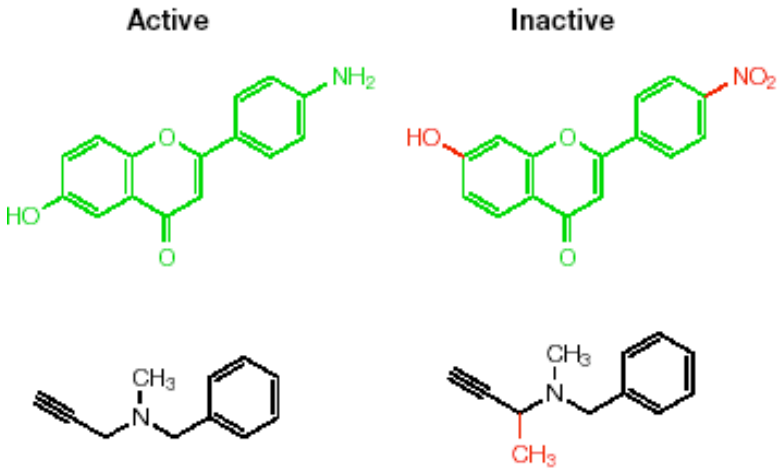
Get a big haystack, find a needle in it

High Throughput Screening

- Implemented in 1990s, still going
- Libraries 1 million compounds in size
- Didn’t live up to hype
  - Single screen program cost ~$75,000
  - Estimated that only 4 small molecules with roots in combinatorial chemistry made it to clinical development by 2001
- Problem: Haystack’s big, but doesn’t have a needle

More Problems

- Can make library even bigger if you spend more, but can't get comprehensive coverage
  - Estimated that 10^{50} to 10^{130} molecules with weight <1000 Da estimated
- Similarity paradox
  - Slight change can mean difference between active and inactive



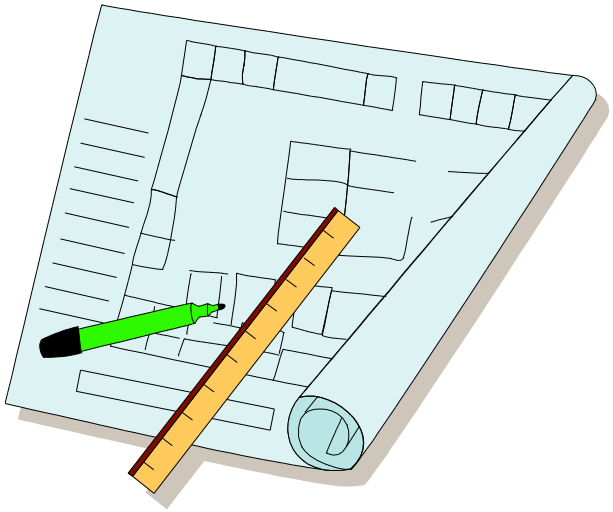
Computation to the Rescue?

- Library design
- Virtual screening
  - Look through library in a computer, much faster/cheaper than experiment
  - Can be used to narrow down candidates for experimental screen
  - Range of methods
    - Drug likeness tests
    - Similarity searches
    - QSAR
    - Docking
    - Free energy computation
  - Can even look beyond binding, to ADMET and drug interactions

3. Design

- Today, “rational” or “structure-based design by a structural biologist or medicinal chemist

- We’ll talk about de novo design



Class Details

Guha. January 10, 2006

Aims

- Solid base of knowledge, whether you go to a big pharmaceutical company, a biotech company, a software startup, or pursue research
- Familiarity with powerful new methods coming online
- Comfort with the literature and discussion that generates new ideas

C.S. Issues, but Applied

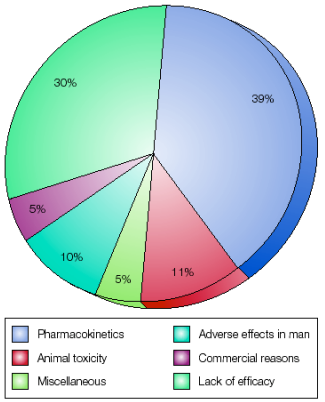
- Searching/sampling high dimensional space
- Machine learning
- Large scale databases
- Geometric algorithms
- Simulation
- Parallelization
- Hardware (clusters, GPUs, specialized boards)

Requirements

- High ratio of material/utility to amount of work
  - Much depends on your effort and interest
  - What work there is will impact whole class
- Every week: read, attend, bring 2 or 3 questions/comments
- Couple weeks: present papers and lead discussion of them
- Final week: brief case study of actual application of computation to drug discovery, or original proposal of a method or application
- Grade breakdown roughly follows time: 30% participation, 60% presentations, 10% case study

Schedule

- Introduction, History, Why Compute
- Search, Pharmacophores, and QSAR
- Docking
- Molecular Mechanics and MM-PBSA
- Free Energy Calculation
- Designing Libraries
- Designing Small Molecules
- In Silico ADME (absorption-distribution-metabolism-excretion)
- Computational Infrastructures
- Case Studies



Web and Email

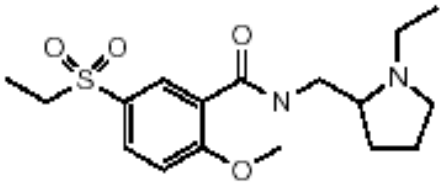
- cs379a.stanford.edu
  - Notes, links to reading, and presentations will be posted
- guha@stanford.edu, Clark S296

Next Week

1D

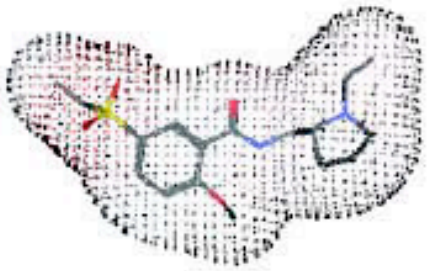
$C_{17}H_{26}N_2O_4S$ → molecular mass

2D



→ number of aromatic bonds; molecular connectivity index; logP(o/w)

3D



→ van der Waals volume; solvent-accessible surface area

Bajorath, 2002

Next Week Continued

- Pharmacophores
  - Specific arrangement of particular features that are thought to give a molecule its activity
  - If you can identify a good pharmacophore, then you can search for other molecules that have it
- QSAR
  - Quantitative structure activity relationship
  - Basically a form of supervised learning

Next Week Readings

- RAPID: Randomized Pharmacophore Identification for Drug Design (Finn, Latombe, Motwani, Yao, et. al.),
- Identification of... Growth Hormone Secretagogue Agonists by Virtual Screening and Structure-Activity Relationship Analysis (J. Med. Chem.),
- QSAR analysis of anticonvulsant agents using *k* nearest neighbor and simulated annealing PLS methods (J. Med. Chem.)

Links up on web, don't get stuck on chemical details, set up proxy if you need off campus access