Informatics and Visualization Tools for Pharmacogenetics Research

Thomas Ferrin, Ph.D.

Resource for Biocomputing, Visualization, and Informatics

University of California, San Francisco

Resource for Biocomputing, Visualization, and Informatics

The RBVI is a NIH/NCRR Biomedical Technology Research Center

We create innovative computational and visualization-based data analysis methods and algorithms, turn these into easy-to-use software tools, and apply these tools for solving a wide range of genomic and molecular recognition problems within the complex sequence → structure → function triad

Application areas

Gene characterization and interpretation

Drug design

Variation in drug response due to genetic factors

Protein engineering

Biomaterials design

Bioremediation

Prediction of protein function from sequence and structure

Sample RVBI projects

- New methods for large-scale data collection, storage, analysis, and presentation for polymorphism (SNP) genotyping project

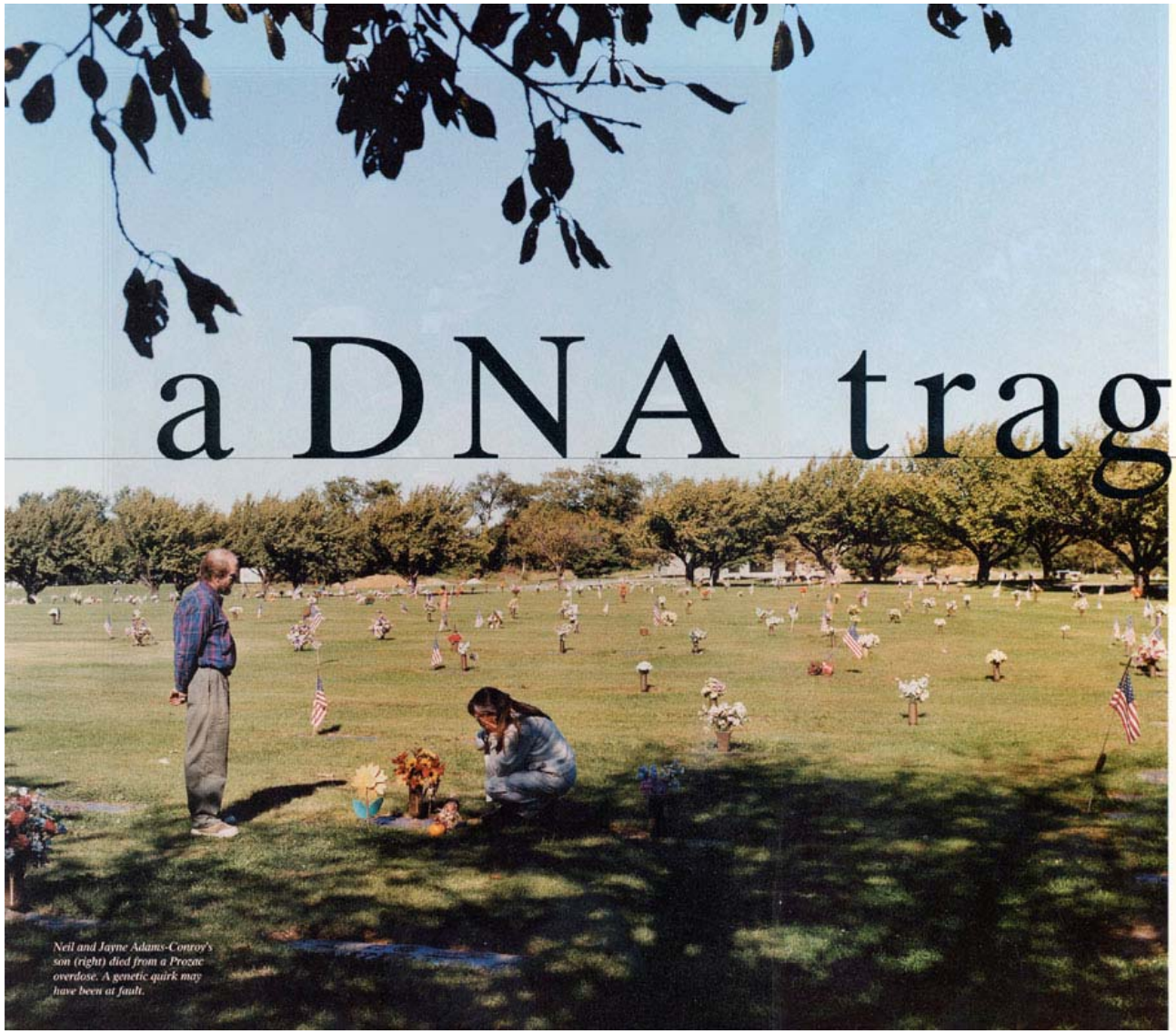
- Extensible visualization tools for comparative studies of protein sequence, structure, and function

a DNA tragedy

Genetic tests to prevent adverse drug reactions may save tens of thousands of lives a year, but for a troubled boy named Michael they came too late.

By David Stipp

Photographs by Suzanne Opton



Neil and Jayne Adams-Conroy's son (right) died from a Prozac overdose. A genetic quirk may have been at fault.

THE DEATH OF NINE-YEAR-OLD MICHAEL ADAMS-CONROY didn't seem at first like a signal event in medicine. It seemed like homicide.

Michael's short life was an uphill struggle from the start. Malnourished as an infant, he was taken from an abusive mother and placed in a temporary foster home before his first birthday. By the time he was 6, his medical record bulged with bad news: Michael was cognitively blunted and violently moody, and appeared to be afflicted with the brain damage of fetal alcohol syndrome, as well as with obsessive-compulsive disorder, tic-inducing Tourette's syndrome, and attention-deficit hyperactivity disorder.

Over the next few years he achieved a semblance of normalcy, thanks to the steadying hands of the resolutely affectionate couple who adopted him at age 3 and to daily doses of drugs to check his tics and obsessions. Small for his age, he took pride at finally being able to fling his coat up onto the grownups' pegs at his home in Martins Creek, Pa., a one-stoplight town two hours north of Philadelphia. He was learning to bowl in a league for handicapped kids and help his dad tend the garden.

Fortune - Oct 30, 2000

Case Report #1: Michael Adams-Conroy

Young child born to abusive mother, adopted at age 3, with signs of fetal alcohol syndrome, obsessive-compulsive disorder, Tourette's syndrome, and attention-deficit hyperactivity disorder. Prescribed Prozac to help control emotional outbursts.

Child dies suddenly; toxicology tests show massive overdose of Prozac. Adoptive parents investigated for homicide and their other two children put into protective custody.

Michael Adams-Conroy (continued)

Sharp-eyed psychiatrist notices unusually high levels of other metabolites in toxicology report, indicating child may have had an enzyme deficiency inhibiting Prozac from being metabolized normally.

Subsequent genetic testing showed child had defect in 2D6 gene which resulted in abnormal liver enzyme that metabolizes antidepressants.

Adoptive parents exonerated.

Case Report #2

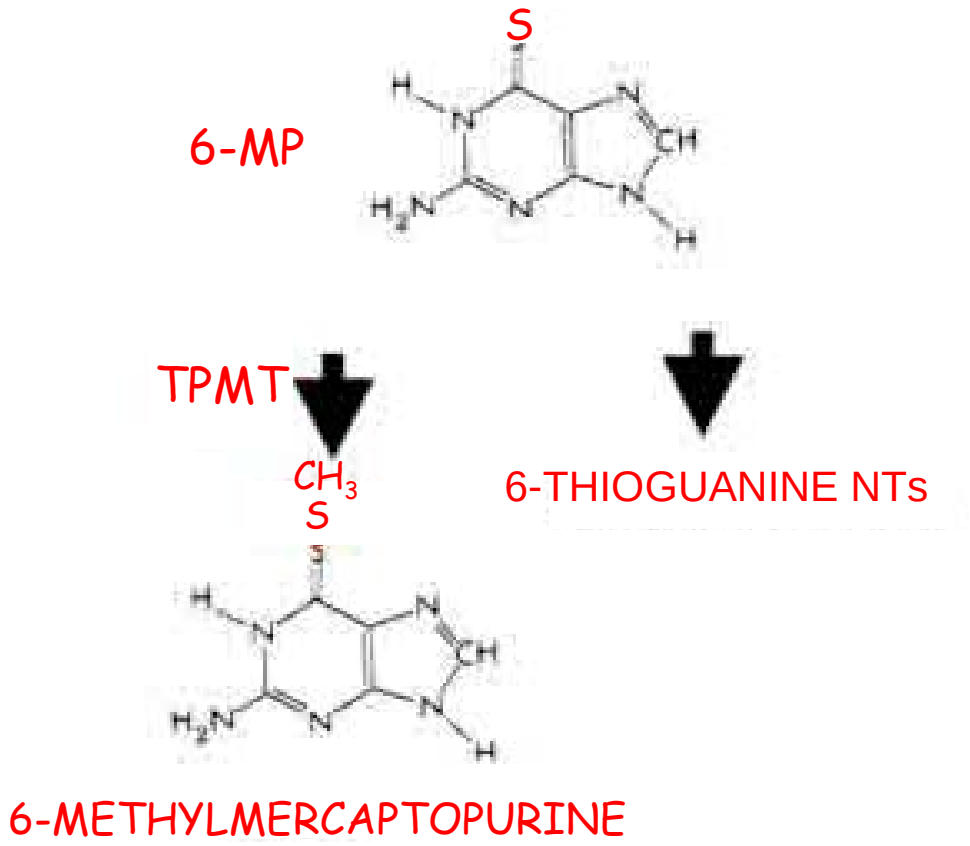
Patient: 3-year old boy

Diagnosis: Acute Lymphoblastic Leukemia (ALL)

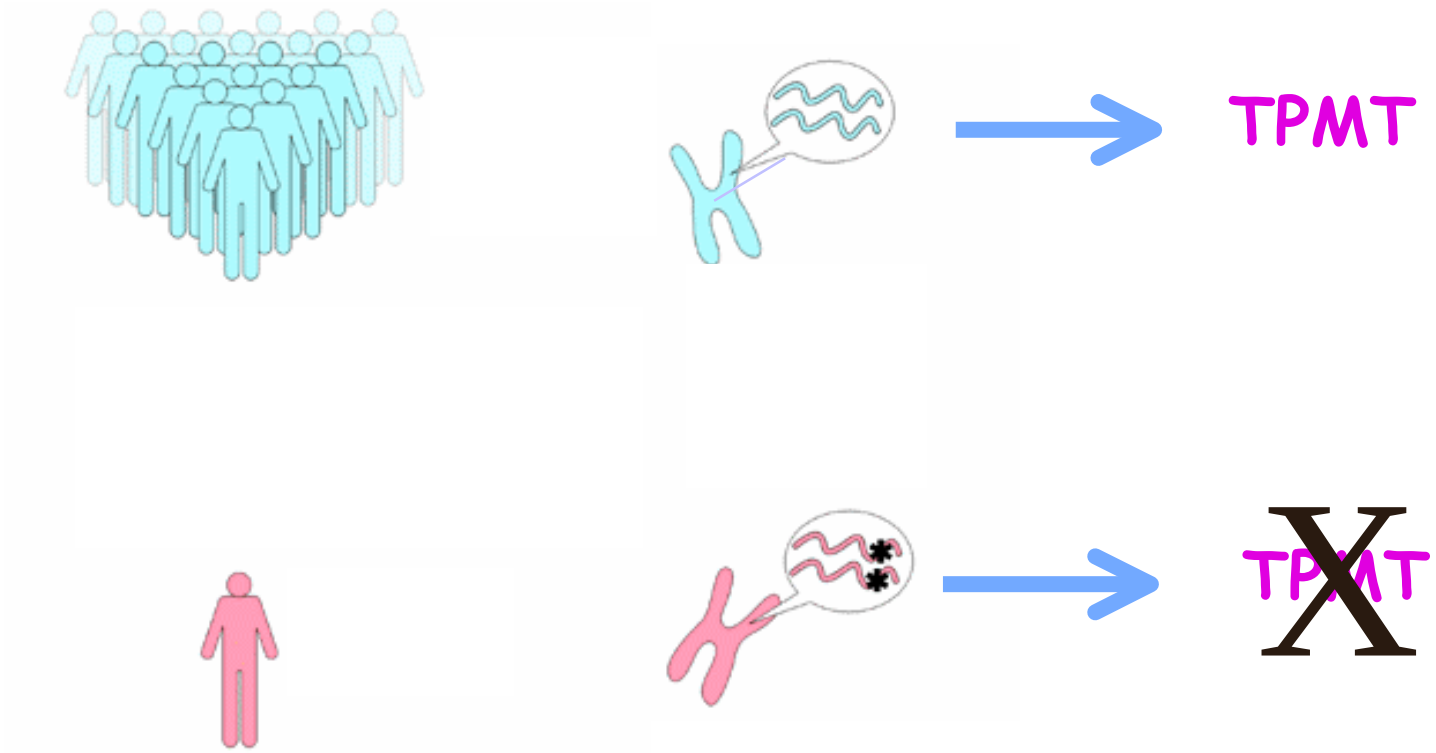
Standard therapy: 6-mercaptopurine (6-MP)

Result: Adverse Drug Reaction leading to acute bone marrow suppression

Normal Mechanism of Action

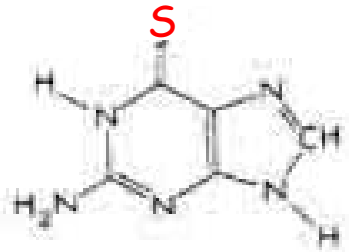


THIOPURINE METHYLTRANSFERASE (TPMT) GENES ARE DEFECTIVE IN 1:300 PEOPLE



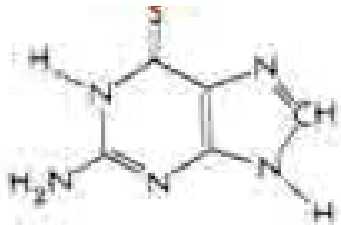
This leads to elevated levels of Thioguanine Nucleotides

6-MP



TPMT

6-THIOGUANINE NTs ↑↑↑



6-METHYLMERCAPTOPURINE

PEOPLE DIFFER IN THEIR RESPONSE TO DRUGS

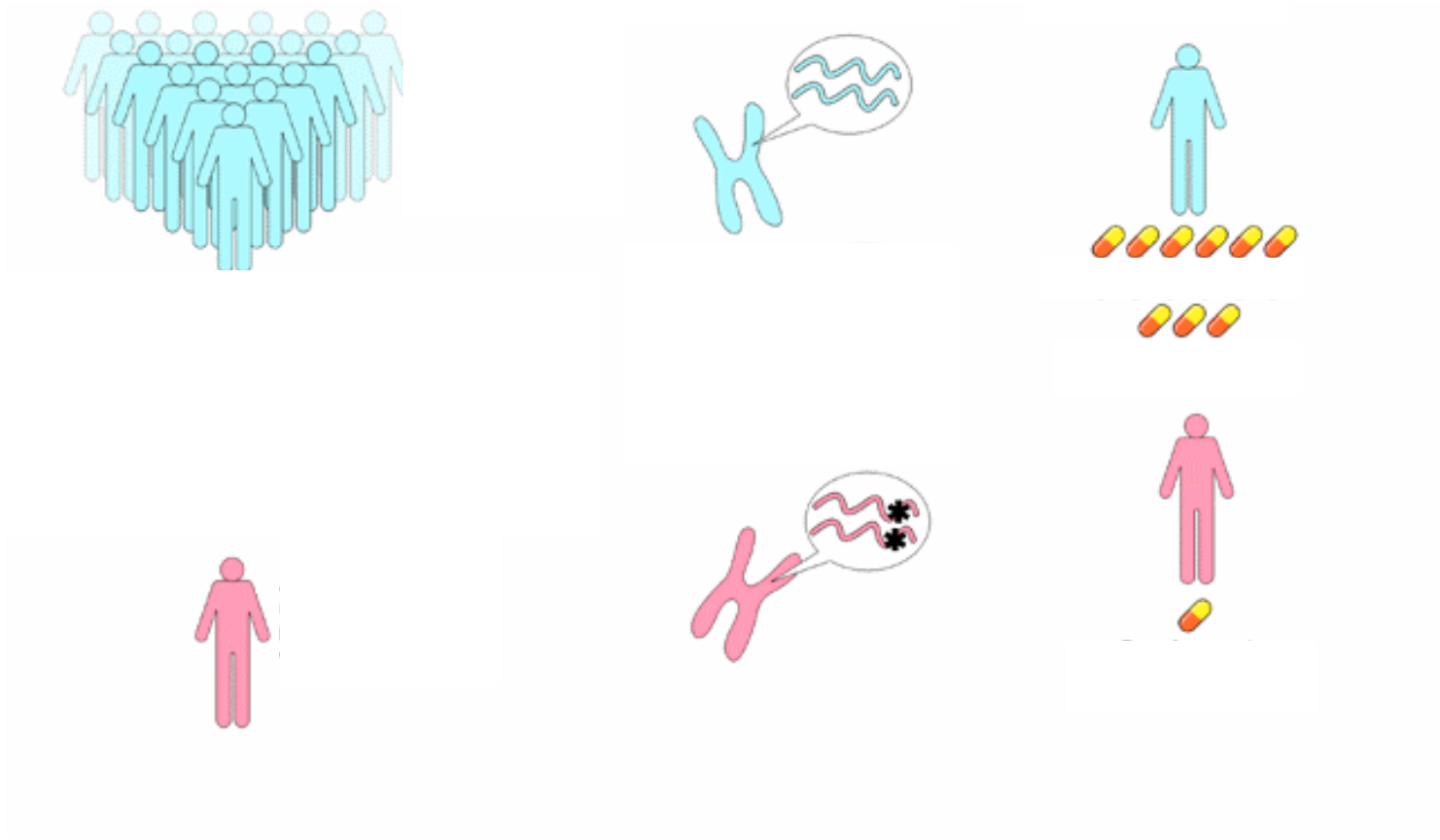
NO RESPONSE

THERAPEUTIC RESPONSE

ADVERSE DRUG REACTION (ADR)

TESTING FOR TPMT GENES IS NOW AVAILABLE

CHILDREN WITH DEFECTIVE TPMT GENES SHOULD RECEIVE A LOWER DOSE OF 6-MP



Adverse Drug Reactions

ADRs may kill 30,000 - 40,000 Americans each year and cause 2,200,000 serious nonfatal reactions. JAMA 1998 June 3;279(21):1684

Drugs with known genetically-linked potential for fatal adverse reactions (partial list):

Drug (Brand Name)	Perscribed For...	Adverse Reaction	Gene at Cause
Imipramine (Tofrannil)	Depression, ATD	Heartbeat irregularity	CYP2D6
Isoniazid (Laniazid)	Tuberculosis	Liver toxicity	NAT2
Warfarin (Coumadin)	Preventation of blood clots	Internal bleeding	CYP2C9
5-fluorouracil (Adrucil)	Cancer	Severe immune suppression	DPD
Clarithromycin (Biaxin)	Antibiotic	Heartbeat irregularity	KCNE2
Azathioprine (Imuran)	Rheumatoid arthritis	Severe immune suppression	TPMT

Pharmacogenetics and Pharmacogenomics

Pharmacogenetics

- The study of the genetic basis for variation in drug response

Pharmacogenomics

- The use of genetics and genomics in drug discovery and development

Research Goals:

- Identify responders and non-responders before therapy
- Predict response to other related drugs
- Avoid toxicity in certain drugs

Potential Impact:

- Major impact on drug discovery process
- Major impact on prediction of drug response
- Avoidance of toxic effects in many individuals

Pharmacogenetics of Membrane Transporters

$12-million, 4-year NIH grant

- Kathleen Giacomini, PI, plus 20 other UCSF researchers

Major Project Goal:

- Understand the genetic basis for variation in response to drugs which interact with membrane transporters. This class of proteins is of great pharmacological importance, as it provides the target for about 30% of the most commonly used prescription drugs and is a major determinant of the absorption, distribution and elimination of many others.

PMT project goals - continued

- Determine the amount of genetic variation (single-nucleotide polymorphisms) in at least 40 transporter genes by examining the DNA from an ethnically diverse sample of 250 people.
- Test the performance of these transporter variants in cell cultures and determine, through clinical phenotype studies, if people with those variants respond differently to drugs in a clinically significant way.
- Provide access to the data from these studies to the general scientific community through the World Wide Web to facilitate collaborative research and to speed development of new drug treatments.

The Corriel Cell Collection

African American (AA) - 100

Caucasian (CA) - 100

Asian American (AS) - 30

Mexican American (ME) - 10

Pacific Islander (PA) - 7

TOTAL - 247

Pharmacogenetics.ucsf.edu - Mozilla {Build ID: 2002031104}

File Edit View Search Go Bookmarks Tasks Help Debug QA

Back Forward Reload Stop http://pharmacogenetics.ucsf.edu/ Search Print

Home Bookmarks FaxReader Google Web Development

Pharmacogenetics.UCSF.edu

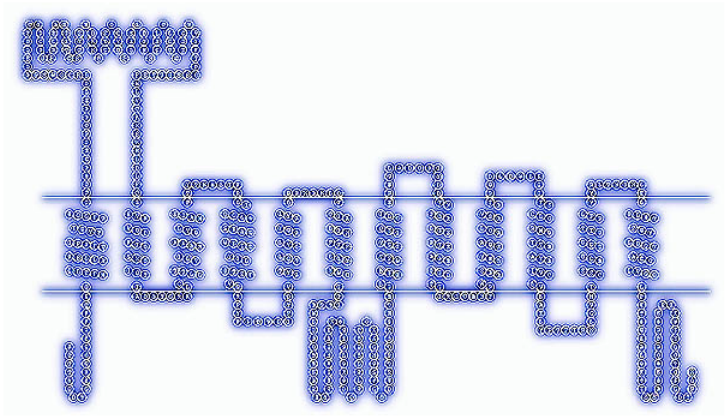
Contents

- home
- goals
- people
- institutions
- publications
- available data
- training
- relevant links
- contact us

pmt intranet
(password protected)

UCSF Pharmacogenetics of Membrane Transporters

The UCSF Pharmacogenetics of Membrane Transporters (PMT) Project is sponsored by the National Institutes of Health's National Institute of General Medical Sciences (grant U01 GM61390). The Project is part of the Pharmacogenetics Research Network and Knowledgebase. Information about the entire Network can be found at PharmGKB. Pharmacogenetics is the study of the genetic basis for variation from person to person in response to drugs. Membrane transporters play a major role in drug response in two ways. First, many drugs work by affecting function of transporters. Second, transporters determine the level of drugs within the body and thus determine whether drug levels are adequately high for therapeutic effect. The goal of the UCSF PMT Project is to understand the genetic basis for variation in drug response for drugs which interact with membrane transporters.



Model of organic cation transporter (OCT2)

Document: Done (0.23 secs)

OCT2 Transporter - Mozilla {Build ID: 2002031104}

File Edit View Search Go Bookmarks Tasks Help Debug QA

Back Forward Reload Stop https://pharmacogenetics.ucsf.edu:8001/caf-results/OCT2/index.html Search Print

Home Bookmarks FaxReader Google Web Development

OCT2 Transporter

PMT Project Investigator

Dr. Kathleen Giacomini

HGNC symbol

SLC22A2

Chromosome

6q26

Aliases

ORGANIC CATION TRANSPORTER 2; OCT2
SOLUTE CARRIER FAMILY 22, MEMBER 2; SLC22A2

Background information

NCBI data

Omim data
LocusLink data

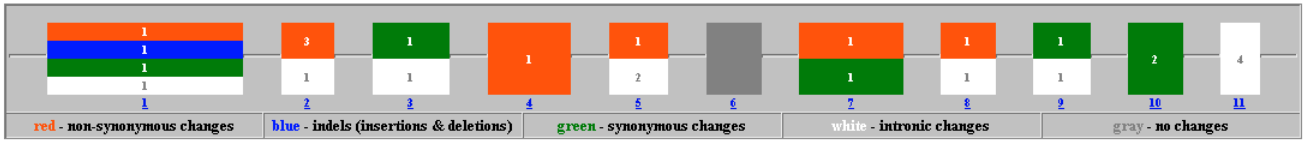
Reference Entry

X98333.1
Homo sapiens mRNA for organic cation transporter, kidney.

Exons

A schematic representation of the gene showing its exons with the exons scaled in size relative to one another. Variant results are indicated by color(s) of the exon block.

Click on number to view desired individual exon results.



Variants

The following results were obtained using 247 ethnically identified DNA samples from the Coriell Institute.

Version 5 of data analysis results, generated on Jun 11, 2002.

Variant identification summary (tab-delimited text version)
Variant identification summary by ethnicity
Variant identification per-sample data (list) (tab-delimited text version)
Variant identification per-sample data (plot)
Population genetics statistics (tab-delimited text version)
Consensus sequences with mammalian species

Document: Done (0.25 secs)

OCT2 Transporter - Mozilla {Build ID: 2002031104}

File Edit View Search Go Bookmarks Tasks Help Debug QA

Back Forward Reload Stop https://pharmacogenetics.ucsf.edu:8001/caf-results/OCT2/index.html Search Print

Home Bookmarks FaxReader Google Web Development

red - non-synonymous changes | blue - indels (insertions & deletions) | green - synonymous changes | white - intronic changes | gray - no changes

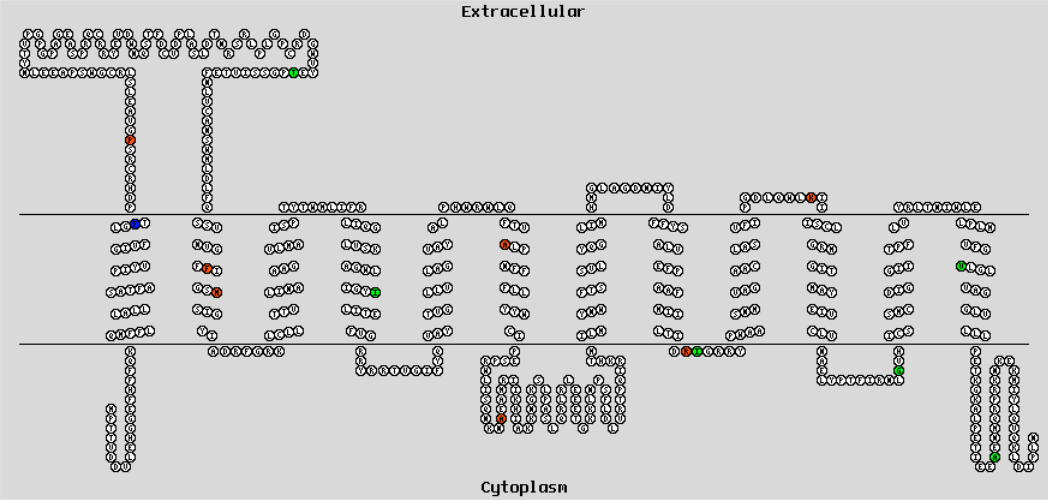
Variants

The following results were obtained using 247 ethnically identified DNA samples from the Coriell Institute.

Version 5 of data analysis results, generated on Jun 11, 2002.

Variant identification summary (tab-delimited text version)
Variant identification summary by ethnicity
Variant identification per-sample data (list) (tab-delimited text version)
Variant identification per-sample data (plot)
Population genetics statistics (tab-delimited text version)
Consensus sequences with mammalian species

Transmembrane prediction for OCT2 protein. Non-synonymous amino acid changes shown in **red**, indels (insertions and deletions) in **blue**, and synonymous changes in **green**.



Cellular Phenotyping Results

No data available yet.

Clinical Studies

No data available yet.

Generated on Tue Jun 11 18:34:12 2002 using software developed by the UCSF Pharmacogenetics of Membrane Transporters project

Document: Done (0.25 secs)

PMT Intranet Website

Used by ~100 researchers at UCSF

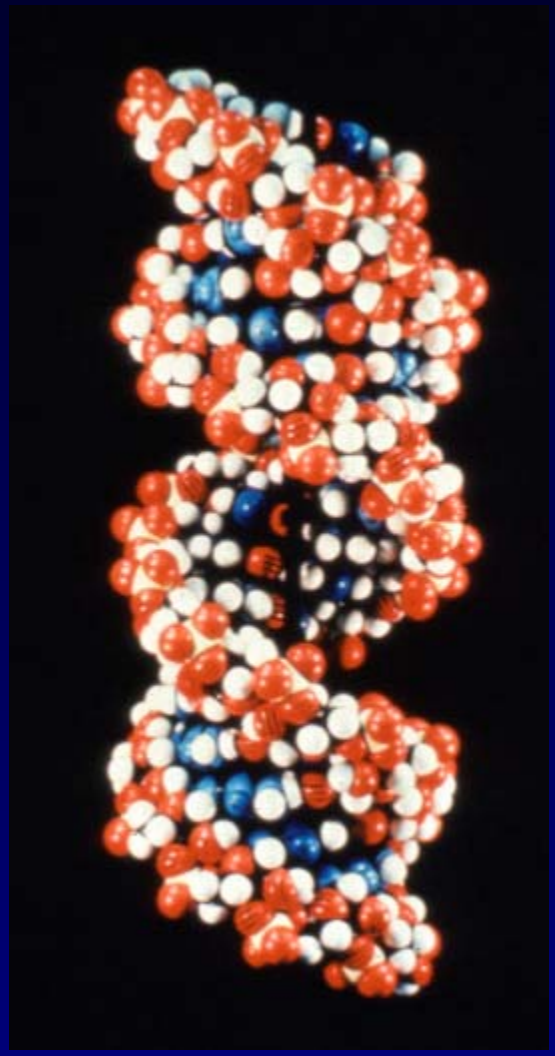
Effective data analysis and display driven by iterative design/refinement cycle, successful because the bioinformatics team works closely with the molecular biologists

- Jill Mesirov, Whitehead: "Bioinformatics needs to be tightly integrated with the scientific research, not a service function"

Flexibility key!

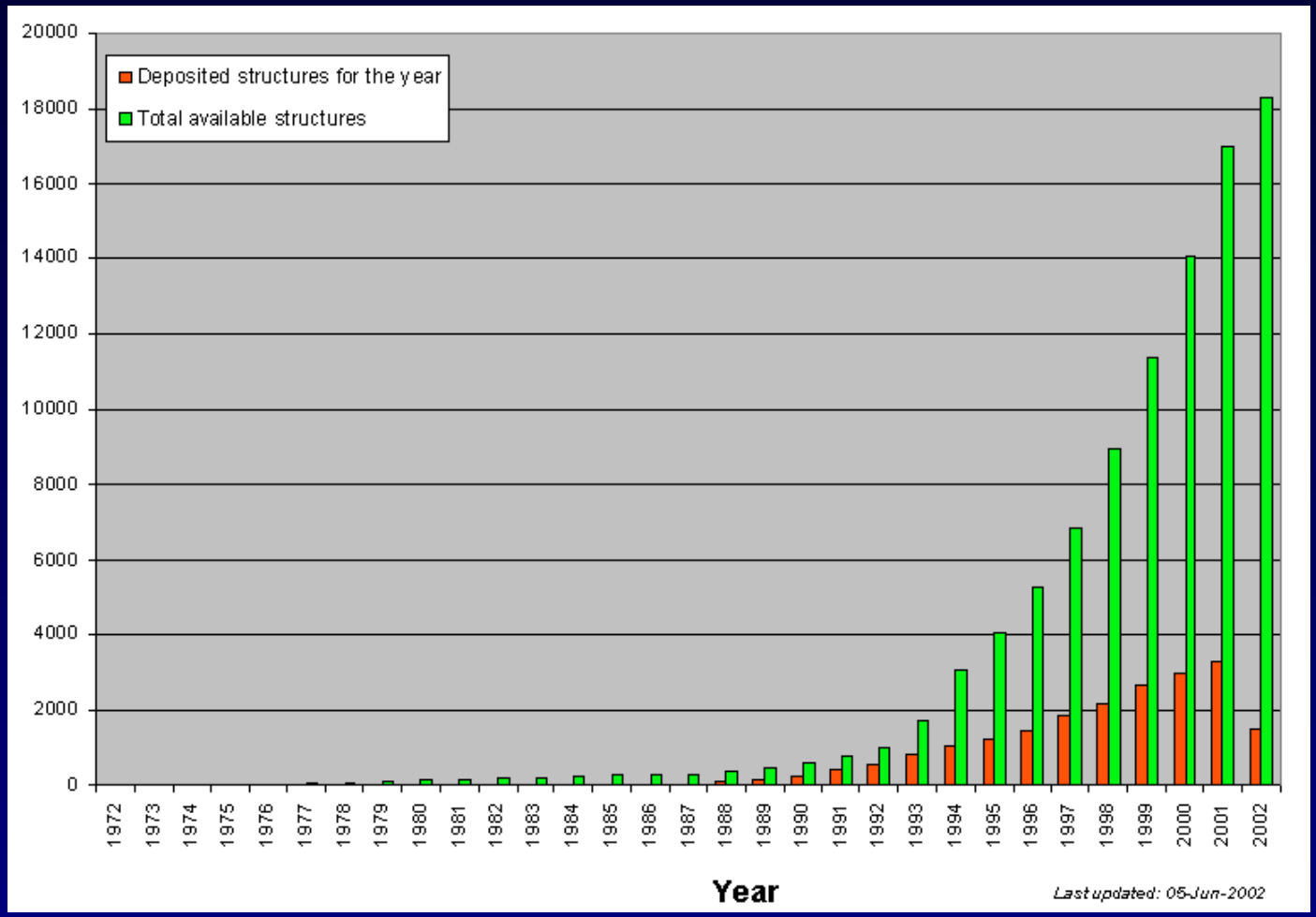
- Multiple ways to display same data
- Simple download mechanism for scientists who want to load raw data into Excel spreadsheets

You've seen sequences.



What about structure?

Growth in Protein Structures



The Structural Genomics Initiatives

"The next step beyond the human genome project"

$150 million in NIH grants to establish 9 U.S. centers

- Goals:
  - Speed the determination of three-dimensional atomic-scale maps of proteins
  - 35,000 structures by 2005
  - Identify all proteins expressed in an organism - "proteomics"

Center	Lead Institution	Target
NY Struct. Genomics Res. Consortium	Rockefeller Univ.	Bacteria/yeast/human
Northeast Struct. Gen. Consort.	Rutgers Univ.	Roundworm/fly/human
Southeast Collab. for Struct. Gen.	Univ. of Georgia	Bacteria/roundworm/human
Berkeley Struct. Genomic Center	Lawrence Berkeley Lab.	Bacteria
Joint Ctr. for Struc. Genomics	Scripps Research Inst.	Roundworm/human
TB Struct. Genomics Consortium	Los Alamos Nat. Lab.	Tuberculosis
Midwest Ctr. for Struct. Genomics	Argonne National Lab.	Archaea/bacteria/eukarya
Ctr. for Eukaryotic Struct. Genomics	Univ. of Wisconsin	*Arabidopsys thaliana*
Struct. Gen. of Pathogenic Protozoa	Univ. of Washington	Protozoans

See http://www.nigms.nih.gov/funding/psi.html for additional information

The X-ray Crystallography Pipeline

Target Selection

"Crystallomics"
- Isolation
- Expression
- Purification
- Crystallization

Data Collection

Structure Solution

Structure Refinement

Functional Annotation

Publish

Anticipated Bioinformatics Developments:
- Determination of distant homologs and domain recognition
- Automation of procedures through empirical rules
- Software integration for decision support and data analysis
- Automated annotation through sequence and structural alignments, prediction of protein-protein and protein-ligand interactions, and motif recognition

Sequence → Structure → Function

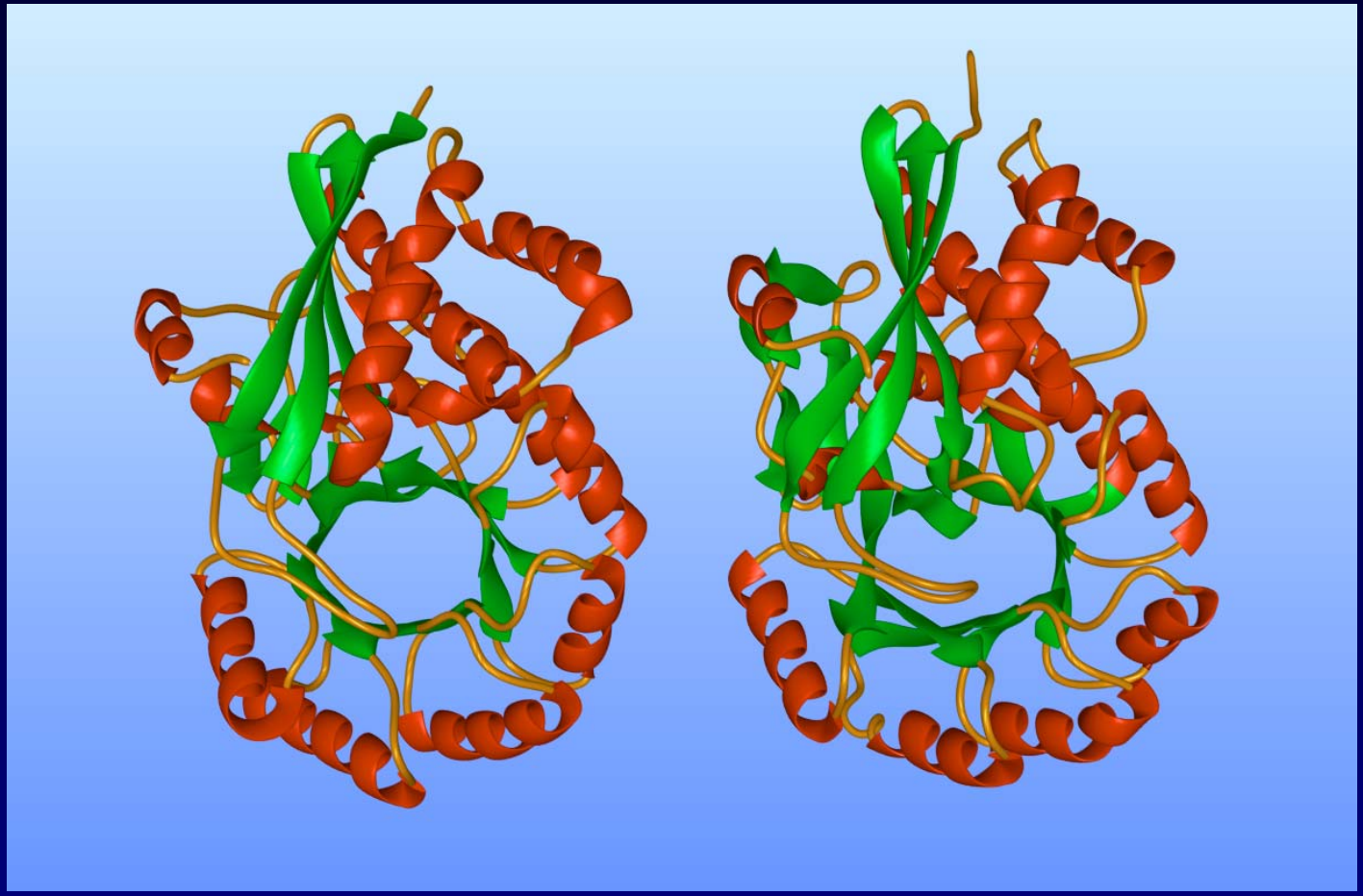
Challenges:

- Prediction of structure from sequence
- Prediction of function from sequence
- Understanding of evolutionary changes
- Engineering proteins for specialized function
- Applications in pharmacogenetics

Potential for major impact on…

- Drug discovery and development
- Prediction of drug response
- Avoidance of toxic side effects

The Same Protein?



Chimera Molecular Modeling System

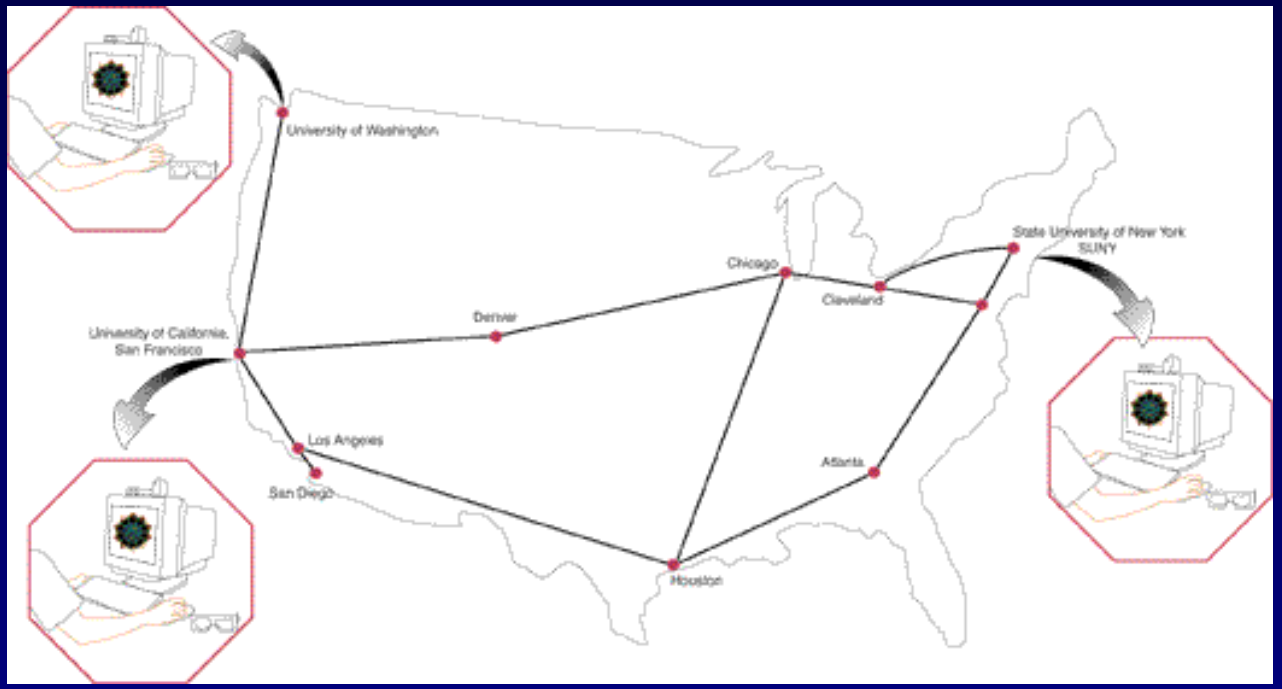
Chimera is an extensible interactive 3-D modeling system designed to allow developers to quickly incorporate novel algorithms and analysis tools

- ~30 extensions written to date
- Extensions are written in the Python programming language
  - Easy to learn, even for novice programmers
  - Offers object-oriented language features
- Extensions can control standard user interface features (e.g. camera, help, menus, toolbar) as well as their own custom interfaces

Sample Chimera Extension

Collaboratory

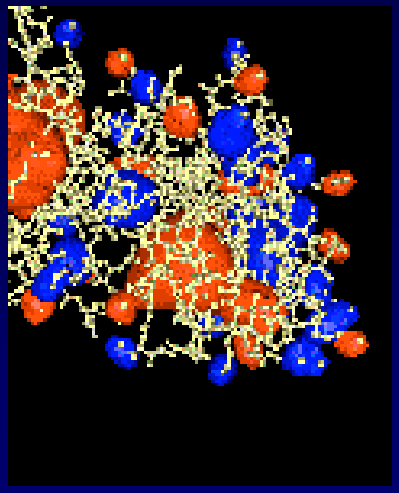
- supports collaborative studies of molecular structure among scientists at multiple remote locations



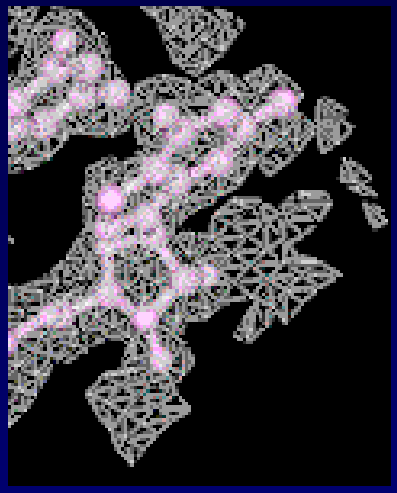
Sample Chimera Extension

Volume Viewer

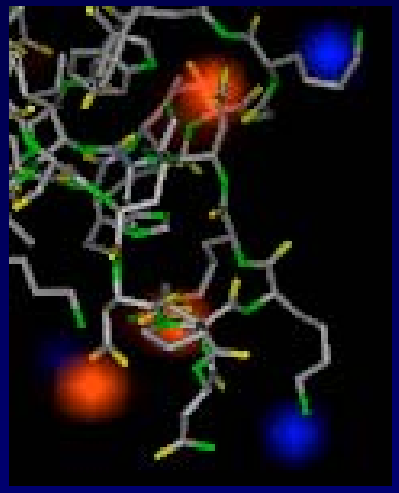
- an extension for visualizing three-dimensional (3D) numerical data sets



Electrostatic potential (surfaces)



Electron density (mesh)

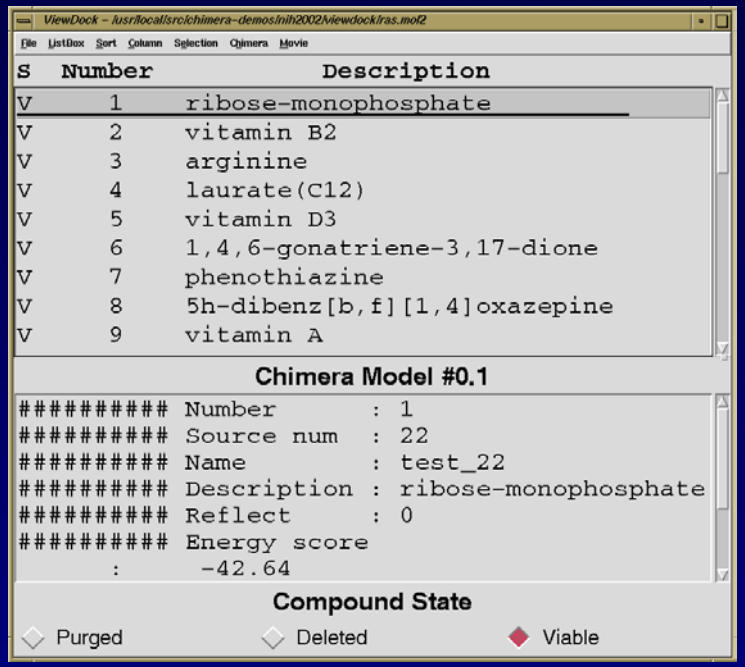


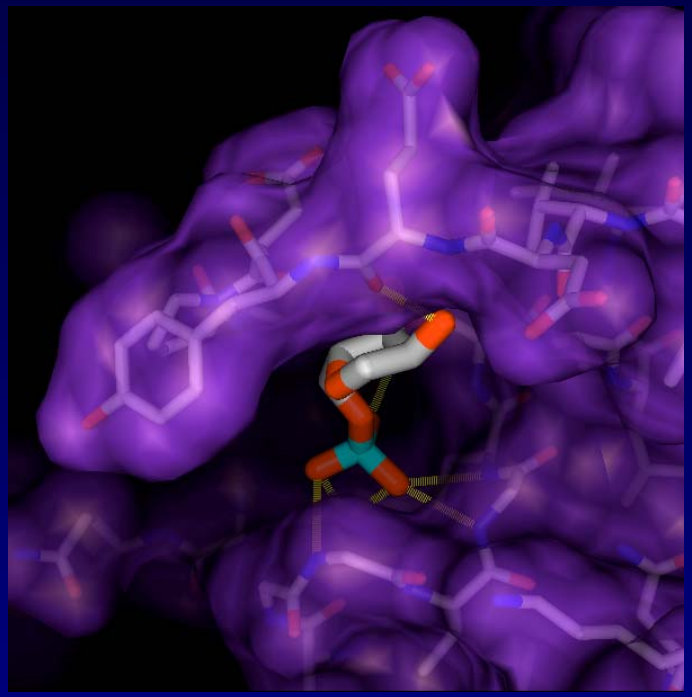
Electrostatic potential (solids)

Sample Chimera Extension

ViewDock

- rapid screening of promising drug candidates found with the UCSF DOCK program

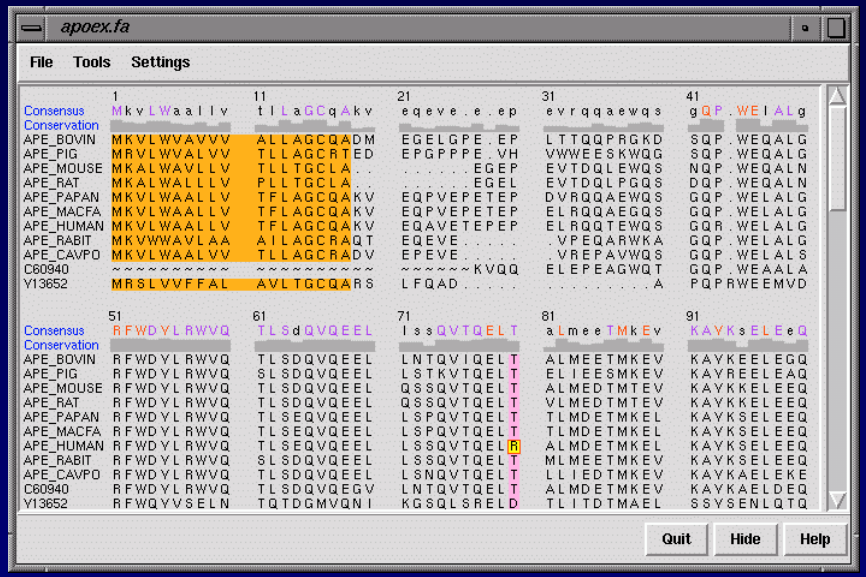




Sample Chimera Extension

Multalign Viewer

- simultaneous display of protein sequence and structure



Chimera Demonstration...

Demo disclaimer:

- Murphy's Law in effect
  - The probability that something will go wrong is directly proportional to number of people in this room

Hardware:

- Dell laptop - 1.2Ghz Pentium 3 w/512MB RAM
- NVIDIA Quadro2 Go 3-D graphics chip

Software:

- Windows 2000
- OpenGL, Python, Tkinter
- Chimera beta version 1

Summary

We are in the midst of a profound and exciting new era in computational biology

The data made available by the various genome and structural genomics projects will occupy researchers for decades to come

High performance computing and the internet play a critical role in the navigation, analysis, and dissemination of this data and the resulting scientific knowledge

The potential impact on drug development and treatment of human disease is enormous

Acknowledgements

Collaborators & Staff

- Dr. Conrad Huang, Dr. Elaine Meng, Prof. Patricia Babbitt, Prof. Kathy Giacomini, Greg Couch, Eric Pettersen, Al Conde, Tom Goddard, Susan Johns, Doug Stryke, Michiko Kawamoto

NIH National Center for Research Resources

- P41-RR01081

National Institute of General Medical Sciences

- GM61390

Additional information

RBVI:
www.cgl.ucsf.edu

PMT project:
www.pharmacogenetics.ucsf.edu

Chimera:
www.cgl.ucsf.edu/chimera