# The Chimera Collaboratory

### **Thomas Ferrin**

### Resource for Biocomputing, Visualization, and Informatics

### University of California, San Francisco



Data and Collaboratories in the Biomedical Community

### Resource for Biocomputing, Visualization, and Informatics

We create innovative computational and visualization-based data analysis methods and algorithms, turns 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  $\rightarrow$  structure  $\rightarrow$  function triad

### Application areas

Gene characterization and interpretation

Drug design

Variation in drug response due to genetic factors

Protein engineering

**Biomaterials design** 

Prediction of protein function from sequence and structure

### Motivation for Collaboratories

#### Science today, especially that which is biology-related, **REQUIRES** collaboration for success

- Multi-institution papers the norm
  Solid evidence of the benefits of "face-to-face" collaborations
- Testimony from researchers that "collaboration promotes new ideas"

#### Desire to extend this productive environment to remote scientists

- Promote more frequent and spontaneous collaborative projects
- Encourage new collaborations

#### Desire to utilize remote collaborative environments in new ways

- Training
- Improved feedback for software development

# **Collaboratory Goals**

Create an interactive three-dimensional graphics application to support collaborative studies of molecular sequence/structure/function among scientists at multiple remote locations for research and training

- Multiple scientists at remote locations able to interactively manipulate shared, complex three-dimensional molecular models (as similar as possible to "face-to-face" collaborations)
- Full semantics for the modification of an object by any collaborative participant
  - access to the underlying molecular data, not just the molecule's graphical representation
  - individual participants can perform operations privately first, then present results in collaborative session
- Small number (n<4) of participants for collaborative research
- Modest number (10-20?) of participants for training

### 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

- Extensions can control standard user interface features (e.g. camera, help, menus, toolbar) as well as their own custom interfaces
- Extensions are written in the Python programming language
  - Python easy to learn, even for novice programmers
  - Python is object-oriented and provides features needed for development of complex codes
  - ~30 extensions written to date





### **Multalign Viewer**

• simultaneous display of protein sequence and structure

🛋 apoex.fa 🛛 🔹						
File Tool	s Settings					
Conservation APE_BOVIN APE_PIG APE_PIG APE_PAGN APE_PAPAN APE_MACFA APE_HUMAN APE_RABIT APE_CAVPO C00940 V10620	1 MKVLWAAUIV MRVLWVAVVV MRALWAVLVV MKALWAVLVV MKVLWAALLV MKVLWAALLV MKVWAVLAA MKVWAALLV MKVWAALAVV	11 t   LagCqAkv ALLAGCQADM TLLAGCRTED TLLTGCLA PLLTGCLA TFLAGCQAKV TFLAGCQAKV TFLAGCQAKV ALLAGCRAQT ALLAGCRAQT ALLAGCRADV	21 eqeve.e.ep EGELGPE.EP EPGPPPE.VH EGEP EQPVEPETEP EQAVETEPE EQAVETEPEP EQAVETEPEP EQAVETEPEP EQAVETEPEP EQAVETEPEP EQEVE EPEVE	31 evrqqaewqs LTTQQPRGKD VWWESKWQG EVTDQLEWQS EVTDQLEWQS EVTDQLAEWQS ELRQQAEQQS ELRQQAEQQS .VPEQARWKA .VREPAVWQS ELEPEAGWQT	41 g Q P WEIALg S Q P WEQALG S Q P WEQALG N Q P WEQALN D Q P WEQALN D Q P WEQALN G Q P WELALG G Q P WELALG G Q P WELALS G Q P WELALS G Q P WEALALS G Q P WEALALS G Q P WEALALS	
Consensus Conservation APE_PIG APE_PIG APE_RAT APE_PAPAN APE_MACFA APE_HUMAN APE_RABIT APE_CAVPO C60940 Y13652	51 RFWD YL RWVQ RFWD YL RWVQ RFWQ YVSELN	61 TLSDQVQEEL SLSDQVQEEL TLSDQVQEEL TLSDQVQEEL TLSDQVQEEL TLSEQVQEEL TLSEQVQEEL TLSDQVQEEL TLSDQVQEEL TLSDQVQEEL TLSDQVQEEL	71 IssQVTQELT LSTKVTQELT QSSQVTQELT QSSQVTQELT LSPQVTQELT LSPQVTQELT LSPQVTQELT LSSQVTQELT LSSQVTQELT LNTQVTQELT LNTQVTQELT KGSQLSRELD	81 ALMEETMKEV ELIEESMKEV ALMEDTMTEV VLMEDTMTEV TLMDETMKEL ALMDETMKEL ALMDETMKEL MLMEETMKEV LLIEDTMKEV ALMDETMKEV ALMDETMKEV ALMDETMKEV	91 KAYK SELEEQ KAYKELEEQ KAYKELEEQ KAYKKELEEQ KAYKSELEEQ KAYKSELEEQ KAYKSELEEQ KAYKSELEEQ KAYKSELEEQ KAYKSELEEQ KAYKSELEEQ KAYKSELEEQ SSYSENLQTQ	V
				Q	uit Hide He	lp



### Volume Viewer

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







Electron density (mesh)



Electrostatic potential (solids)

### ViewDock

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

	ViewDock - lusrilocalIsrcichimera-demosinih2002Niewdocktras.mol2						
Elle	ListBox Sort Column	Selection Chimera Movie					
ន	Number	Desci	ription				
v	1	ribose-monophe	osphate 🔷				
v	2	vitamin B2					
v	3	arginine					
v	4	laurate(C12)					
v	5	vitamin D3					
v	6	1,4,6-gonatriene-3,17-dione					
v	7	phenothiazine					
v	8	5h-dibenz[b,f]	][1,4]oxazepine				
v	9	vitamin A	7				
Chimera Model #0.1							
##	****	Number :	1				
##	* # # # # # # # #	Source num :	22				
##	****	Name :	test_22				
##	* # # # # # # # #	Description :	ribose-monophosphate				
##	* # # # # # # # #	Reflect :	0				
##	+####### :	Energy score -42.64	7				
	Compound State						
$\diamond$	Purged	Oeleted	🔶 Viable				



### Collaboratory

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



### **Chimera Collaboratory Features**

Commands input on one workstation (typed keywords or menu selections), are displayed and executed on all participating workstations

• Molecular models are displayed on all workstations simultaneously in response to input from any participant

#### Model control provided for each participant

- Interactive control of rotations, translations, scaling, ...
- Colored "telepointers" used by participants to indicate particular regions of interest
- "Private window" for testing ideas
- Participants can join and leave an on-line session at any time

# **Technologies** Employed

#### Chimera

Collaboratory implemented as an extension to the basic system

#### CORBA

• Common Object Request Broker Agent - used to construct rendezvous service for sessions

- COTS desktop videoconferencing
  Prefer standards-based tools (e.g. Polycom's ViaVideo)
  High quality audio imperative, low frame rate video just fine

# Low-latency / modest-bandwidth networking • Initial session startup requires data be replicated remotely, but

otherwise only modest bandwidth required

### Status

Two-participant version largely complete and in active use by our collaborators

Synchronization mechanism being generalized to work with other Chimera extensions

Refinement of user interface on-going

Scaling number of participants just got underway · Social interaction issues may complicate this

Formal evaluation not yet begun

### **Continuing Challenges**

Lack of good quality desktop VTC for Macintosh Session setup (Chimera + VTC) still awkward Difficult to troubleshoot where network bottlenecks occur (e.g. high latency) and what to do about them

### Acknowledgements



#### Collaborators & Staff:

 Dr. Conrad Huang, Dr. David Konerding, Prof. Patricia Babbitt, Prof. John Gerlt (Univ. of Illinois), Dan Greenblatt, Greg Couch, Eric Pettersen, Al Conde, Tom Goddard, Dr. Elaine Meng

Funding: National Center for Research Resources · P41-RR01081

### Further information:

www.cgl.ucsf.edu