Extending Chimera for collaborative molecular visualization

T.E.Ferrin, C. Huang, T.E.Klein

Computer Graphics Laboratory
Department of Pharmaceutical Chemistry
University of California, San Francisco
San Francisco, California  94143- 0446

http://www.cgl.ucsf.edu
To develop a collaboratory environment for carrying out interactive three-dimensional molecular modeling studies

* multiple scientists at remote locations to interactively manipulate shared, complex three-dimensional molecular models ('face-to-face')

* full semantics for the modification of an object by any collaborative participant
  - access to the object's data, not just the object's graphical representation
    - individual participants can perform operations privately first, then present results in collaborative session
typical scenario

* type command on keyword on one workstation, display and execute the command on all other participating on-line collaboration workstations

* display molecules moves in tandem in real-time on all workstations simultaneously in response to input from any participant

* provide independent control for each participant for a separately shaped or colored mouse cursor
  - highlight interesting facets of a molecular model
  - interactively control rotations, translations and scaling

* participants can join and/or leave an on-line session at any time
Computer graphics

- qualitative
- generates pictures
- quality vs. real-time, interaction
- value lies not in numbers themselves, but insights gained
- idea generator

Experimental techniques

- X-ray crystallography
- NMR
- Mass spectroscopy
Thematic Example: Structure-Based Drug Design

* Facile and comprehensive system for determining the structures of proteins and nucleic acids in solution and designing new ligands and drugs

* Integrated Software Tools for Structure-Based Drug Design Applications
Extending Chimera for Collaborative Molecular Visualization

* data exchange among programs
  - standardized data definitions
  - common I/O routines
    - CHIMERA (C++ and Python)
    - AMBER (FORTRAN)
Extending Chimera for Collaborative Molecular Visualization

Goals

Overview

Tools

Collaboratory

Technologies

Research

* Test Cases

- Molecular Mechanisms of Mutagenesis and DNA Repair: Recognition of Damaged DNA

- Dihydrofolate Reductase

- Structural Aspects of type 1 Collagen in Osteogenesis imperfecta
Training

* active learning for local and remote users
  - classroom-style teaching
  - collaboratory-style teaching
* 1-1
* 1-N
* Chimera
  - users local to UCSF
  - users at remote sites with high-speed network access
  - users with today's typical Internet access
Desktop Videoconferencing

* real-time, multi-party multimedia application
* high quality audio
  - better or equal to clear, static free telephone connection
* be able to discern important facial or hand gestures
  - full-motion video
  - reduced frame video
* off-the-shelf
Data Network

* Parameters affecting network performance
  - bandwidth
  - latency
  - delay variance
  - connection quality
  - multicast support
  - quality-versus-price tradeoffs

* Collaboratory Data connection characteristics
  - lower bandwidth than videoconferencing
  - higher reliability
  - bursts of activity
Extending Chimera for Collaborative Molecular Visualization

Overview

- CENIC/CalREN-2 - high speed backbone network infrastructure serving higher education in California
  - map of CalREN-2
  - state-of-the-art data communications
    - minimum internode network bandwidth OC-12 (622 Mbps)
    - UCSF connected via a dedicated OC-3 (155 Mbps) "Packet-over-Sonet" network interface via Cisco router
* CENIC/CalREN-2 connected to vBNS via multiple OC-12 connections
  - map of vBNS
Example of another Collaboratory

* Molecular Interactive Collaborative Environment (MICE)
  - stores molecular scenes in a relational database and queried
  - rendered in VRML
  - does not allow for modification of an object
  - for further information, see http://mice.sdsc.edu
Modeling semantics verses graphics only

Utilize widely available software packages

* TK (graphical interface)
* OpenGL (three-dimensional graphics)
* Python (command language)
  - syntax is straightforward
    - facilitates use by non-programmers
  - embodies object-oriented features
  - extensive libraries available for handling strings, sets, files and graphics
  - source and binaries freely available
  - Copyright allows for free use, even commercial and resale
  - See http://www.python.org
Chimera Software Architecture

Core Functionality

- 3D Graphics (OpenInventor)
- User Interface (Motif, Python)

Data Manipulation Operations (Python, C++)

Python Interpreter (C)

Molecular and Graphics Data Management (OTF, C++)

Extensions

- User Interface (Tk, Motif)
- Extension Functionality (Python, C++, C, Fortran)
Core Functionality

- data management
- user interface and methods of user interaction
- three-dimensional interactive graphics
  - geometric representations of molecules (wireframe, spheres, ball-and-stick, ribbon)
  - molecular surfaces (wireframe, polygonal mesh)
  - text mapping
    - coupling of geometric representations and graphical properties (color and translucency)
  - volume rendering (three-dimensional fields)
    - protein solvent density
    - ensemble probability distributions
    - occupancy distributions from molecular dynamics trajectories
    - isosurface generation
Extending Chimera for Collaborative Molecular Visualization

Chimera

- interpreted command language
- infrastructure support
- hypertext Help system
- single user system

Extensions

- written in Python, C, C++, Fortran, etc.
- built on top of the core functionality
  - World Wide Web capabilities
    - GRAIL (Python web browser)
      - standard web-browsing capabilities
      - download and execute Python (similar to Netscape and Java)
- provide graphical user interface (GUI) for user extensions to basic menu-driven interface
- interaction communications protocol (run CHIMERA on several workstations simultaneously)
Prototype in MidasPlus

* Proof of concept
  * requires identical setup (both software and user data)
* 1-1
* Uses custom protocol over network connection
* Only commands are transmitted (no bulk data)

Development in Chimera

* Solutions for 1-1, 1-N and N-N
  * may be different due to different requirements (e.g. reliability, speed, etc.)
* Communication among participating software clients is central to collaboratory design
* Different technical solutions satisfy different requirements
Solutions under consideration:

- CORBA
  - Common Object Request Broker Architecture
  - From Object Management Group (OMG)
    - Consortium of vendors and end users
  - Distributed objects (similar to remote procedure calls [RPC])
  - Very good for 1-to-1
  - See http://www.corba.org

- Multicast
  - Many implementations (lots of research articles)
  - Most famous is MBONE
  - Good for 1-N where reliability requirements is low
  - Efficient use of network bandwidth

* Target 1-1 first
Designed a real-time, interactive molecular modeling collaboratory to be used for research and training.

Currently being implemented as an extension to the new molecular modeling software Chimera.

See http://cgl.ucsf.edu for updates and availability.
Acknowledgements

People

Thomas E. Ferrin (PI)
Al Conde
Greg Couch
Heidi Houtkooper
Conrad Huang
Teri Klein
Eric Pettersen

Funding

This work is financially supported by the National Institutes of Health (P41-RR-0181)