

Structural Biology on the Grid

SBGrid Research Coordination Network

Harvard Medical School

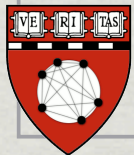
Summary

SBGrid Background
Grid Outreach
Portal Infrastructure
Molecular Replacement
Objectives and Priorities

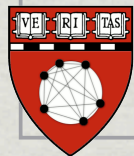


SBGrid

- 📌 Consortium of structural biology labs
 - Involved in structure determination of (primarily) proteins
 - X-Ray crystallography, NMR, Electron Microscopy
- 📌 87 member labs across the US
 - 28 at Harvard & Boston Academic/Medical hub
- 📌 90 software packages
- 📌 Modest local cluster resource
 - 300 cores across several clusters (Intel, Mac, AMD)
- 📌 Now developing web-based portal interfaces to key apps



The SBGrid Team



Ian Stokes-Rees, <http://sbgrid.org>



Structural Biology Grid



[Introduction](#) | [Affiliated Labs](#) | [Software](#) | [Grid Computing](#) | [OS X Computing](#) | [Wiki](#) | [Contact](#)

Welcome to SBGrid

[News](#) | [Services](#) | [History & Staff](#)



Structural Biology Grid (SBGrid) is a computing collaboration of several X-ray crystallography, NMR and electron microscopy laboratories. [Participating laboratories](#) include groups primarily at Harvard Medical School, Harvard University and Yale Medical School, but our alumni often remain members after becoming principal investigators at other institutions.

Explore SBGrid.org

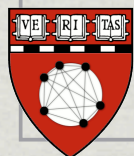
- [See what SBGrid can do for you](#)
- [See the labs currently affiliated with SBGrid](#)
- [Browse the software that SBGrid configures for its affiliates](#)
- [Learn more about the benefits of Grid Computing](#)
- [See how SBGrid is leveraging the power of OS X for scientific applications](#)
- [Download publically-available software released by SBGrid](#)

Latest News

- [SBGrid User Summit: Quo Vadis Structural Biology?](#)
SAVE THE DATE! The first ever SBGrid user summit 'Quo Vadis Structural Biology?' will take place on May 5th and 6th 2008 in Boston, Massachusetts. We will have a number of talks and workshops focused ([read more...](#))
- [Grid computing specialist joins SBGrid](#)
- [Experimental cluster acquired under NSF grant](#)
- [Apple Worldwide Developers Conference](#)
- [New Consortium Members](#)
- [SBGrid Receives Taplin Funds for Discovery Award for Grid Computing](#)

© 2005, Structural Biology Grid

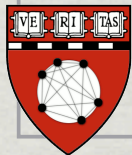
[Member Access](#)



Ian Stokes-Rees, <http://sbgrid.org>

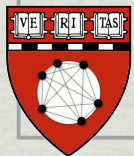
SBGrid Services

- Application optimization
- Application packaging and automated distribution/
update
- Help desk
- Centralized license management
- Soon:
 - Grid portal for SB applications
 - Gateway to OSG



Motivation for Grid

- **Because computational requirements continue to be a bottle neck**
- **Because complexity of tools impedes quality and efficiency of scientific investigation**
- **Because some affiliated labs don't have large compute clusters available to them**
- **Because new computationally intensive methods are being developed**

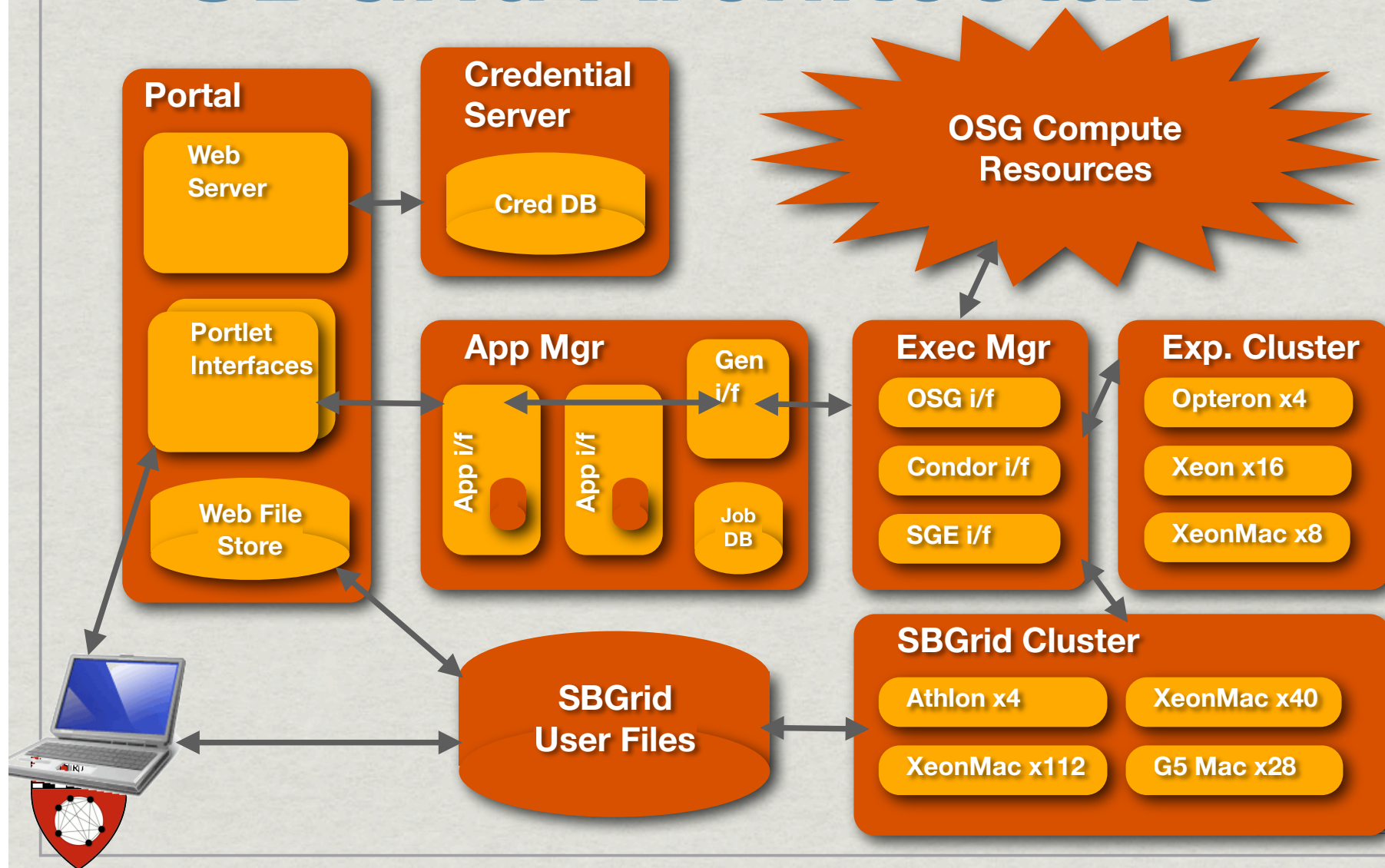


Portal Infrastructure

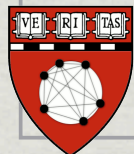
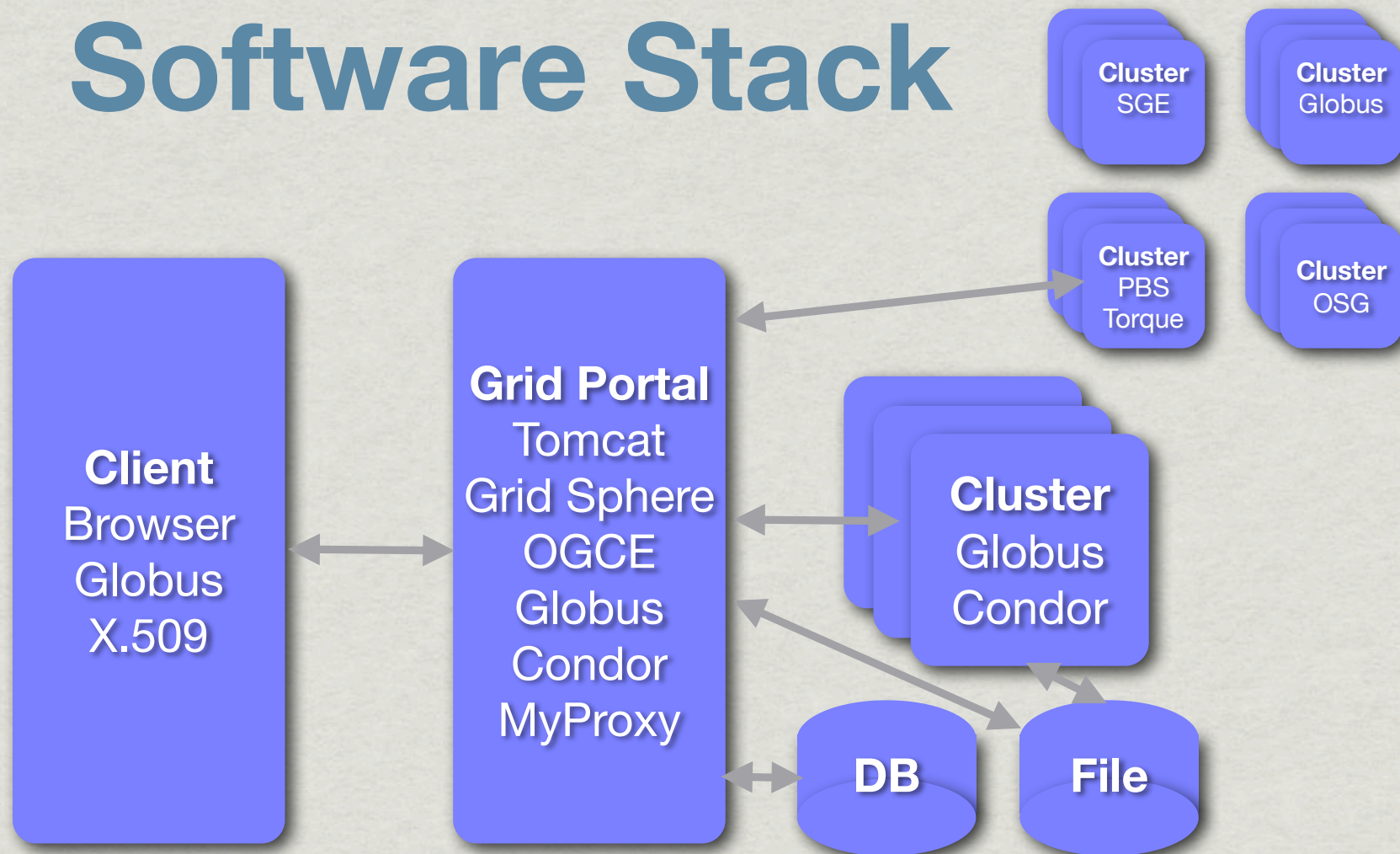


Ian Stokes-Rees, <http://sbgrid.org>

SBGrid Architecture

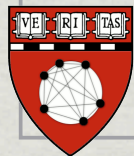


Software Stack



Observations

- Fairly deep software stack
- Development and debugging difficult
- Documentation of existing components often sketchy
- New domains: hard to predict what users want or need



Grid Outreach



Ian Stokes-Rees, <http://sbgrid.org>

Bringing New Users to the Grid

SBGrid consortium

- VO administrators for existing community
- Setting up new users with OSG accounts
- Portal infrastructure to access grid resources and expose “Bio” applications
- Eventually aim to bring member lab computing resources into OSG

Northeastern University

- Partnering with bio-informatics group to introduce OSG
- Providing 6 month internship for undergraduates to deploy/expand SBGrid

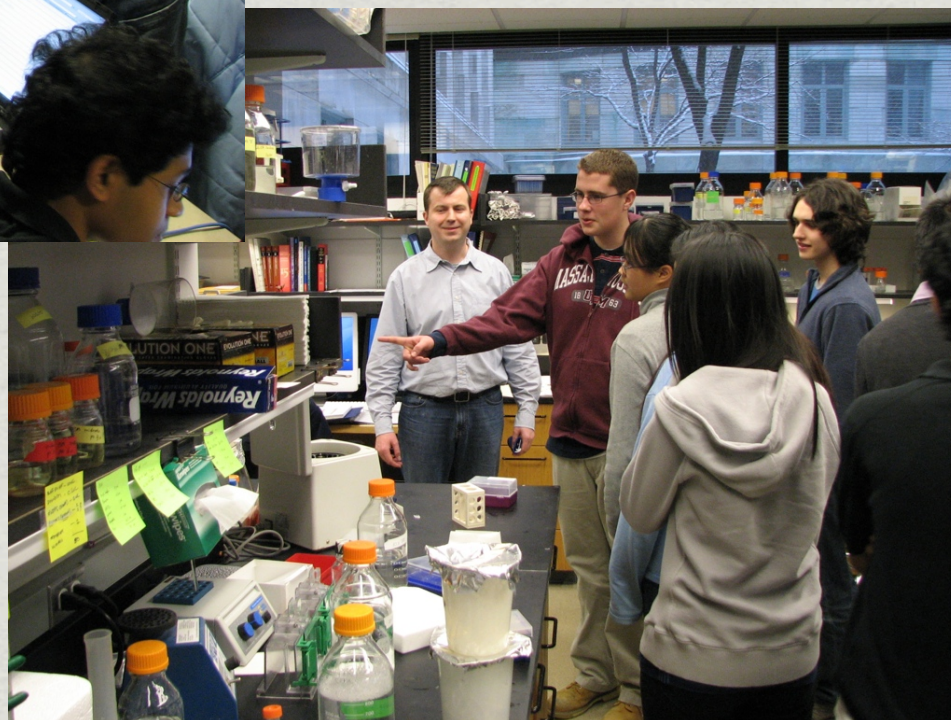
Boston Latin School

- Public exam school, one of the top in US
- Running computational biology seminars and job shadowing
- Summer grid computing internship





Boston Latin School half day computational structural biology seminar



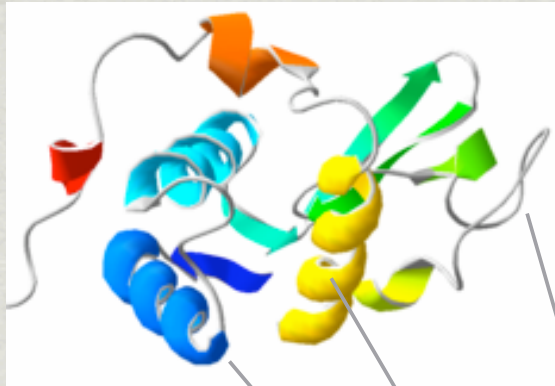
Ian Stokes-Rees, <http://sbgrid.org>



Molecular Replacement



Fourier Transform:

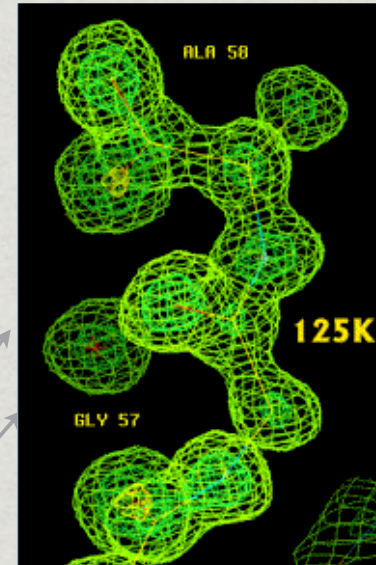


Real
Space

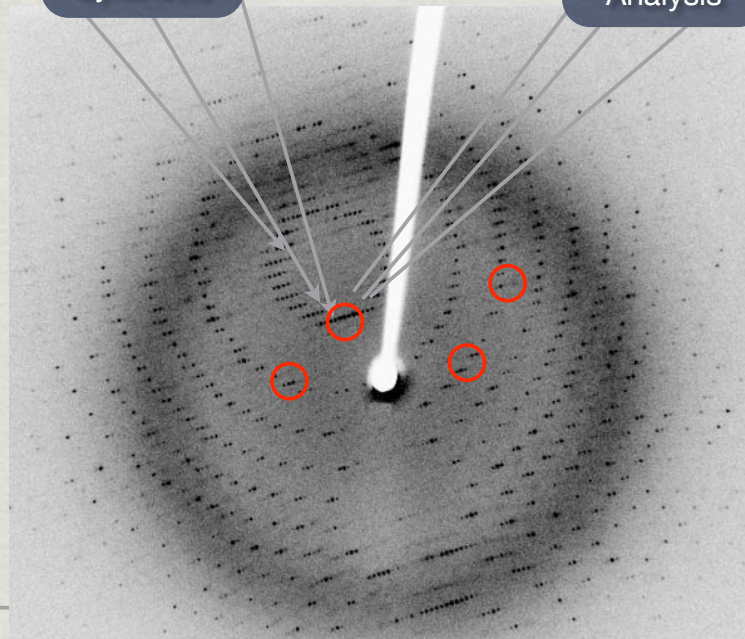
Simple
diffraction
waves

Fourier
Synthesis

Fourier
Analysis

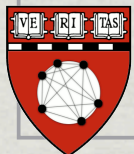


Reciprocal
Space



Fourier series for
electron density is a
sum of contributions
from individual
reflections.

Ian Stokes-Rees, <http://sbgrid.org>



Phase Problem

F_{HKL}

Amplitudes

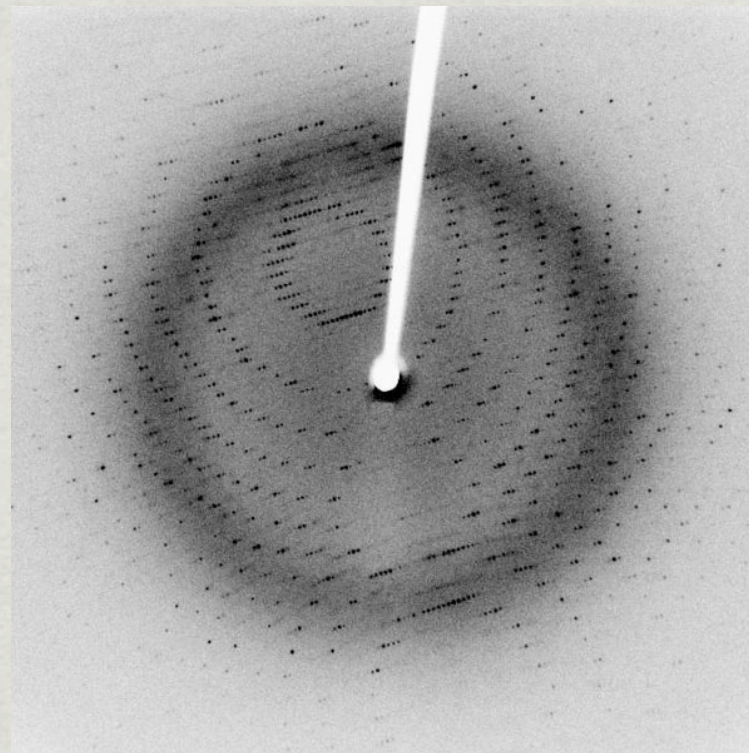
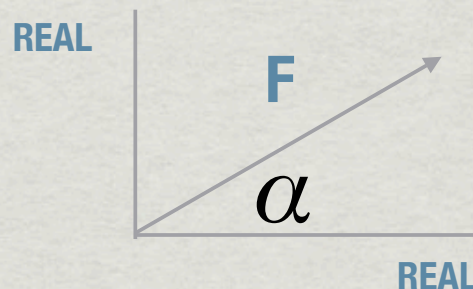
- can be measured
- $\sim \sqrt{\text{intensity}}$

Frequency

- Fixed and known from X-ray source

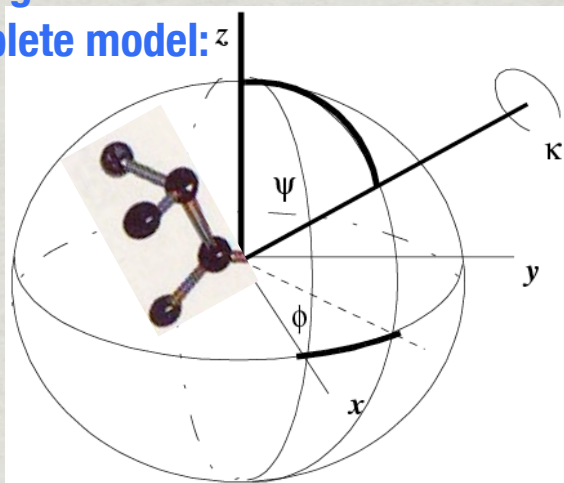
Phase

- Unknown!

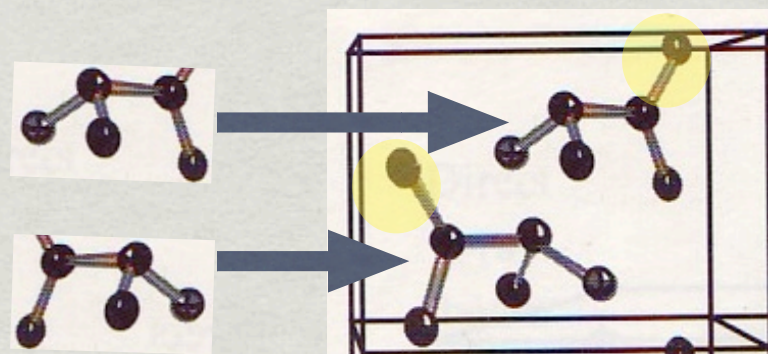


Molecular Replacement

Homologous or
incomplete model:



Rotational Alignment



Translational Alignment



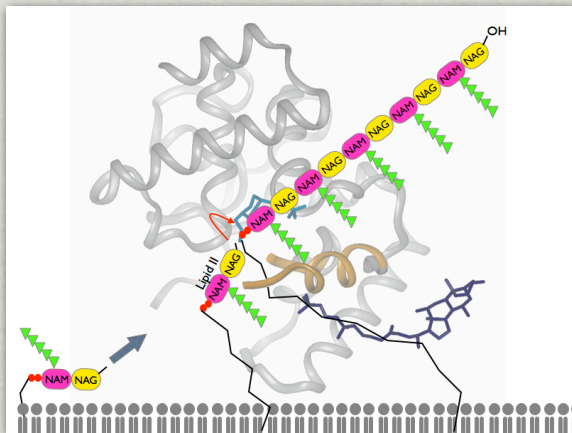
Combining model phases with experimental intensities will reveal details of missing elements

Typically 30% identity and 1/3 of a structure required.

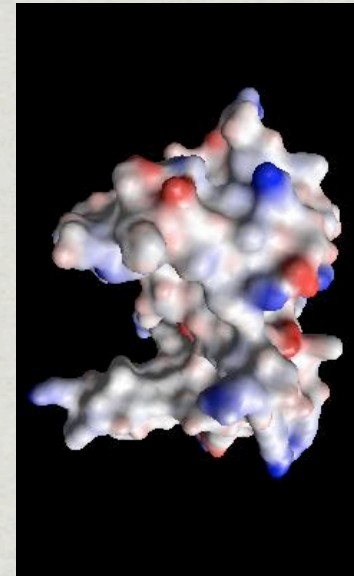


Experimental Phasing

Transglycosylase

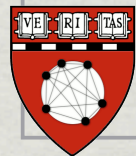


Lysozyme



PHAGE

3 months **Too Slow!**



Quick Tips:



Visit mmcif.rcsb.org for detailed information about the macromolecular Crystallographic Information File (mmCIF) data dictionary.

Are you missing data updates? The PDB archive has moved to [ftp://ftp.wwpdb.org](http://ftp.wwpdb.org). For more information click [here](#).


Welcome to the RCSB PDB

The **RCSB** PDB provides a variety of tools and resources for studying the structures of biological macromolecules and their relationships to sequence, function, and disease.

The RCSB is a member of the **wwPDB** whose mission is to ensure that the PDB archive remains an international resource with uniform data.

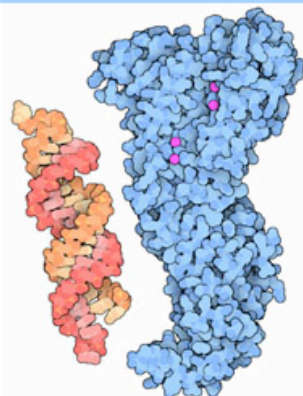
This site offers tools for browsing, searching, and reporting that utilize the data resulting from ongoing efforts to create a more consistent and comprehensive archive.

Information about compatible browsers can be found [here](#).

A **narrated tutorial**  illustrates how to search, navigate, browse, generate reports and visualize structures using this new site. [This requires the Macromedia [Flash player](#) download.]

Comments? info@rcsb.org

Molecule of the Month: Small Interfering RNA (siRNA)



Double-stranded RNA is often a sign of trouble. Our transfer RNA and ribosomes do contain little hairpins that are double-stranded, but most of the free forms of RNA, messenger RNA molecules in particular, are single strands. Many viruses, however, form long stretches of double-stranded RNA as they replicate their genomes. When our cells find double-stranded RNA, it is often a sign of an infection, and they mount a vigorous response that often leads to death of the entire cell. However, plant and animal cells also have a more targeted defense that attacks the viral RNA directly, termed RNA interference.

- [More ...](#)
- [Previous Features](#)

News

- [Complete News](#)
- [Newsletter](#)
- [Discussion Forum](#)
- [Job Listings](#)

19-February-2008



Protein Sculptures on Display at Rutgers

Sculptures and photographs by [Julian Voss-Andreae](#) are currently on display at Rutgers Student Center (New Brunswick, NJ) until February 22, 2008.

- [Full article ...](#)



Winter 2008 RCSB PDB Newsletter Redesigned and Published

Molecular Replacement Structures

PDB ID or keyword Author **MOLECULAR REPLACEMENT** Site Search | Advanced Search

Are you missing data updates? The PDB archive has moved to <ftp://ftp.wwpdb.org>. For more information click [here](#).

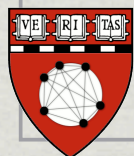
Help 21583 Structure Hits 9953 Citations 4602 Ligand Hits 20 Web Page Hits

Advanced Keyword Query for: molecular replacement
1 2 3 4 5 .. 2159

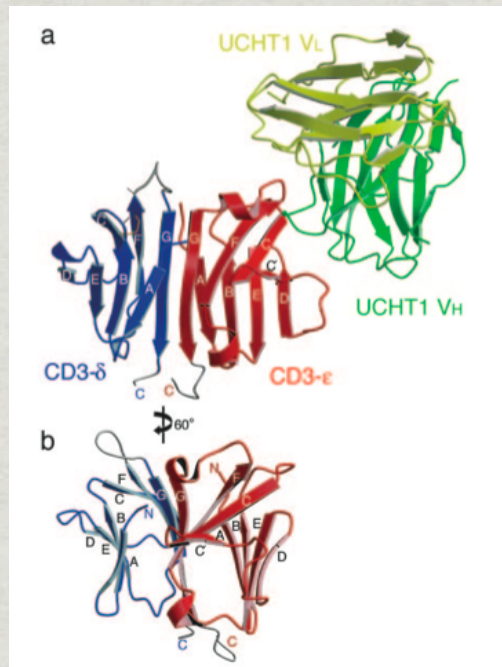
☒ 1m6r    **Crystal structure of rGd(CGCGCG) forming hexamer Z-DNA duplex with 5'-(rG) overhang**

 **Characteristics** **Release Date:** 25-Feb-2003 **Exp. Method:** X Ray Diffraction
Resolution: 1.54 Å
Classification **DNA**
Compound **Polymer:** 1 **Molecule:** 5'-R(*G)D(*CP*GP*CP*GP*CP*G)-3' **Chains:** A,B
Authors **an, B., Sundaralingam, M.**

21583



Example from Harrison Lab, Harvard Medical School



complexes per asymmetric unit. We determined the structure by molecular replacement, implemented with MOLREP (32), using 244 antibody fragments as search models. We used omit maps to eliminate incorrect solutions and to verify the correct one (Protein Data Bank entry 6FAB). We built an initial model with

Our roadmap:



Expand the Antibody Library to incorporate new structures



Setup computations through a portal



Configure molecular replacement applications with more advanced options (e.g. rigid body refinement).

Arnett et al. Crystal structure of a human CD3-epsilon/delta dimer in complex with a UCHT1 single-chain antibody fragment. Proc Natl Acad Sci USA (2004) vol. 101 (46) pp. 16268-73



CASE 2:

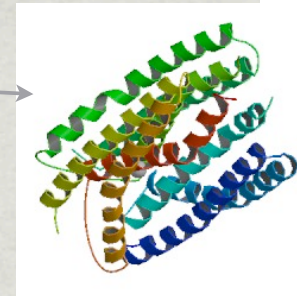
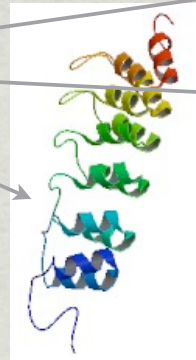
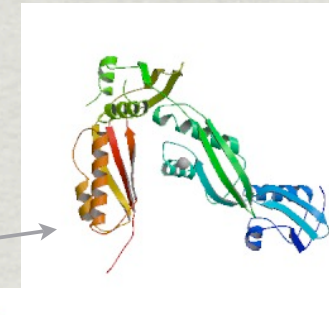
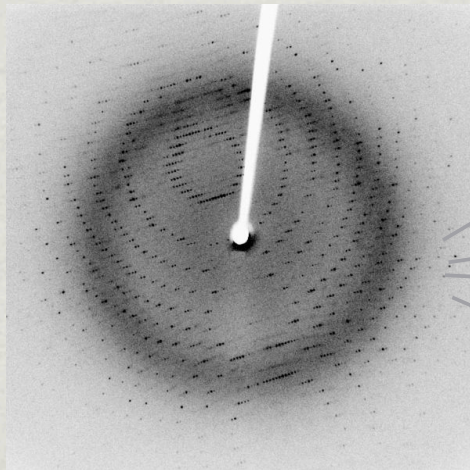
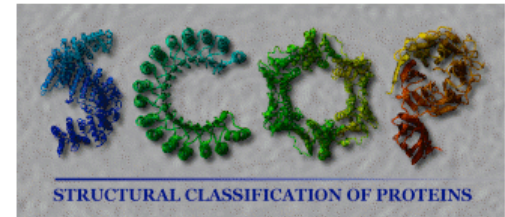
Blind Molecular Replacement

Structural Classification of Proteins

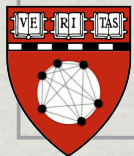


Welcome to **SCOP**: Structural Classification of Proteins.
1.73 release (November 2007)

34494 PDB Entries. 1 Literature Reference. 97178 Domains. (excluding nucleic acids and theoretical models).
Folds, superfamilies, and families [statistics here](#).
[New folds superfamilies families](#).
[List of obsolete entries and their replacements](#).



Ian Stokes-Rees, <http://sbgrid.org>

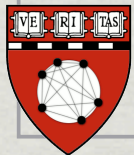


Objectives and Priorities



Grid Computing for Biologists

- * **Ease of use is number one concern**
 - * **Portal infrastructure**
 - * **Single Sign On**
 - * **Single point of access**
- * **Integration of diverse resources**
 - * **Local VDT/OSG-based clusters**
 - * **Local SGE clusters**
 - * **OSG**
 - * **Member labs' compute resources**
- * **Secure processes and data**
- * **Data availability**
 - * **WebDAV**
 - * **SCP/SFTP**
 - * **HTTP(S)**
- * **Storage management**
 - * **Meta-data facilities**
 - * **File catalogue**
- * **Advanced Users**
 - * **APIs**
 - * **Scripting**
 - * **Workflows**



Thank you! Questions?

Ian Stokes-Rees, Research Associate
SBGrid, Harvard Medical School
ijstokes@crystal.harvard.edu

<http://sbgrid.org>

Check out our website and email us
with any questions.