



... for a brighter future

Structural Biology Center: Program Overview and Future Development

Andrzej Joachimiak

APS Cross-cut Review Structural Biology

Argonne, January 24, 2007



U.S. Department
of Energy

UChicago ►
Argonne_{LLC}

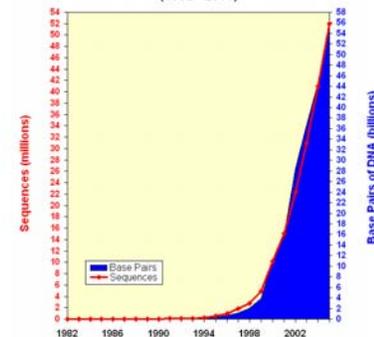


Structural Biology Center Mission

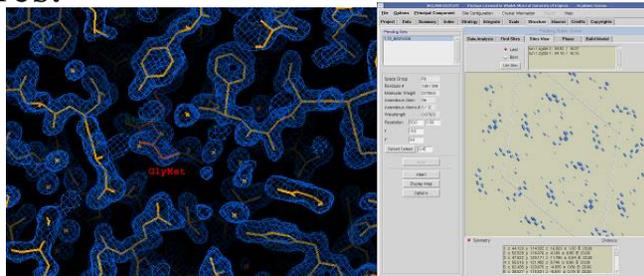
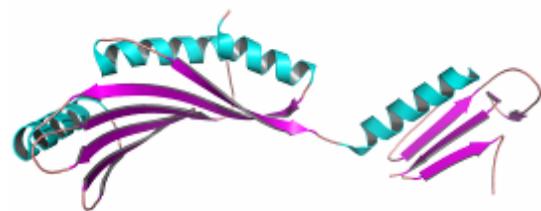
- The SBC mission is to advance and promote scientific and technological innovation in support of the DOE mission by providing world-class scientific research and advancing scientific knowledge
- SBC provides advanced data collection facilities at the APS to the national user community and maintains a high-profile user program
- SBC exploits major advances in macromolecular x-ray crystallography and addresses the most challenging structural biology problems to advance scientific knowledge
- SBC is an important component of integrated biosciences and contributes to the expansion of existing programs and exploration of new opportunities in structural biology, proteomics and genomics research with a major focus on medicine, bio-nanomachines, and biocatalysis that are highly relevant to energy resources, health, a clean environment and national security

Major Changes in Structural Biology

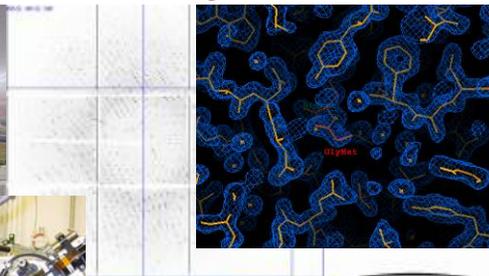
Growth of GenBank
(1982 - 2005)



- We are in a process of conversion structural biology from “cottage industry” to “large-scale science” – SBC is a part of these changes.
- What factors contributed to these changes?
 - Rapid advances in genome sequencing – increased the number of targets
 - Maturation of molecular biology and proteomics technologies – made more samples available for structural studies
 - Development dedicated synchrotron facilities with insertion devices – increase in beam flux and brilliance (small crystals, large assemblies, membrane proteins)
 - Cryo-crystallography – effectively reduced radiation damage
 - Phasing using anomalous signal – *in vivo* incorporation of seleno-methionine into proteins
 - Advances in computing – reduce time to analyze data, determine and refine structures.



Midwest Center for Structural Genomics



Target APC24878

PDB 1J20 F03Eran

Protein: BC2969 (111 aa) 3562126 related

Database crossref: UniProt

Deposited: 09 Jul 2004

Authors: T. Xia, L. Lovrich, A. Jostanek

Target APC24919

PDB 1J16 F0E8an

Protein: BC2354 (235 aa) related proteins

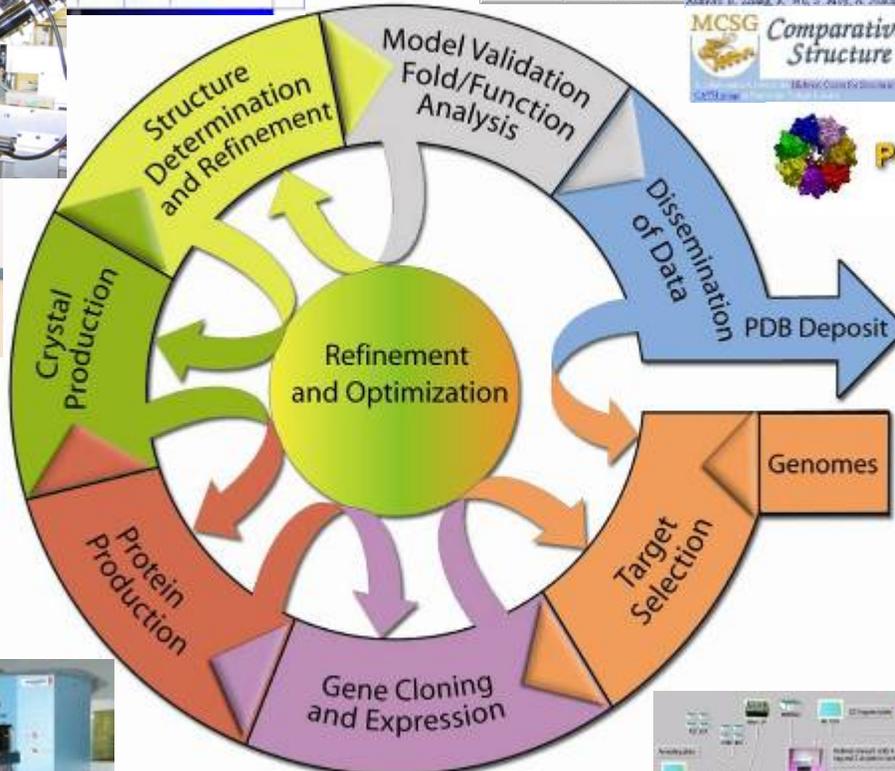
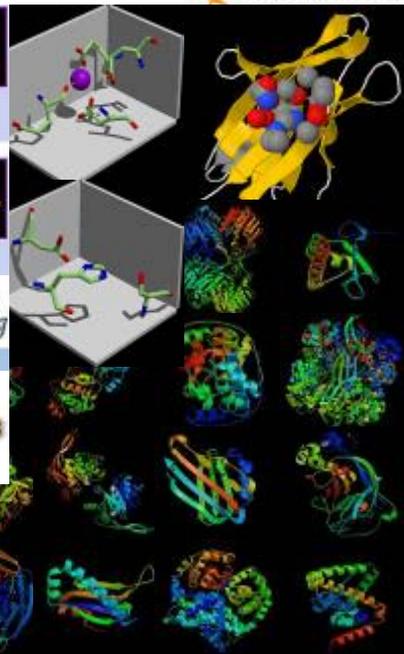
Database crossref: UniProt

Deposited: 07 Apr 2004

Authors: B. Zhang, X. Wu, S. Mir, A. Jostanek

MCSG Comparative Protein Structure Modeling

ProFunc



ATGAGAATGAAGCGATTTTTAA
 TTTGGTTACAAAATTGTTCAAA
 AAATTACAACAAAATTTCAAA
 TTTGAATCACTTGAAAAAATAA
 TCATTAAATTATCCAGTACTAC
 GAGCGAGAACATCGACGTAAAG

APC25529	1E3T1061	Active
APC25531	1E3T1011	Active
APC25570	1E3T1276	Stopped - Gene toxic
APC25534	1E3T1017	Active
APC25536	1E3T1278	Active
APC25540	1E3T1270	Active
APC25541	1E3T1108	Stopped - Gene toxic
APC25544	1E3T1108	Active
APC25545	1E3T1108	Active
APC25547	1E3T1102	Active

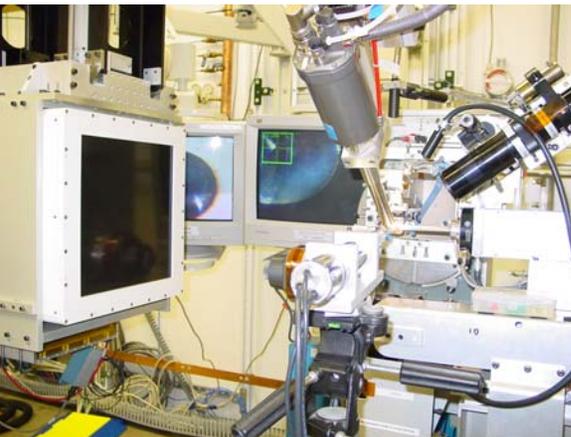
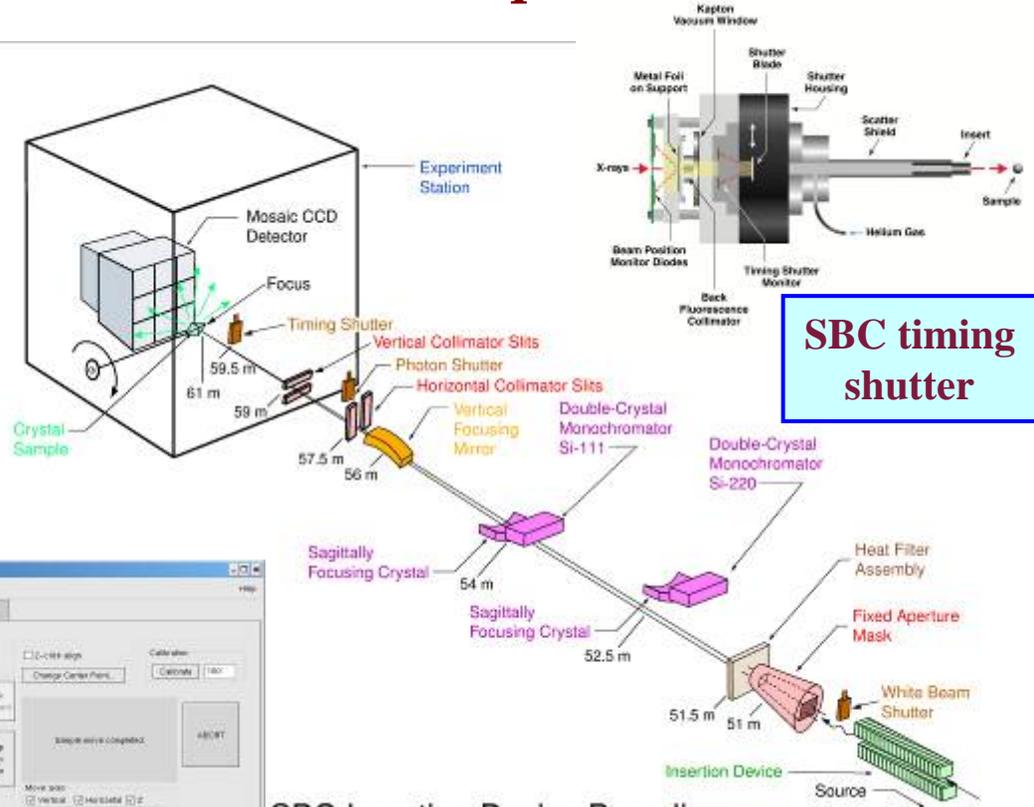
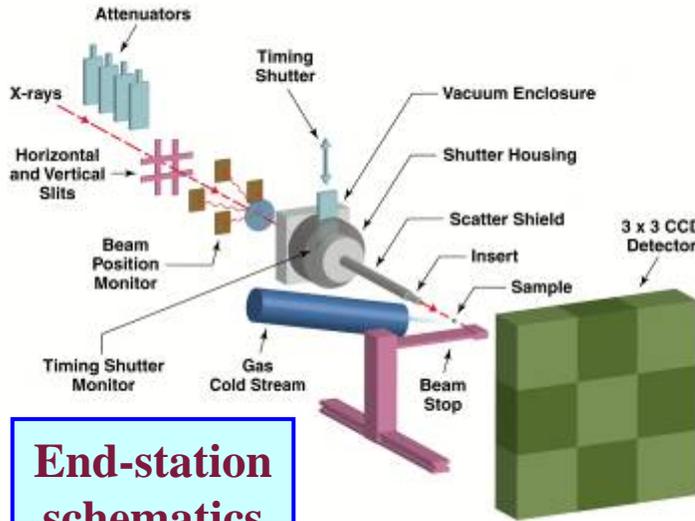


MCSG TaSel Target Selection for MCSG

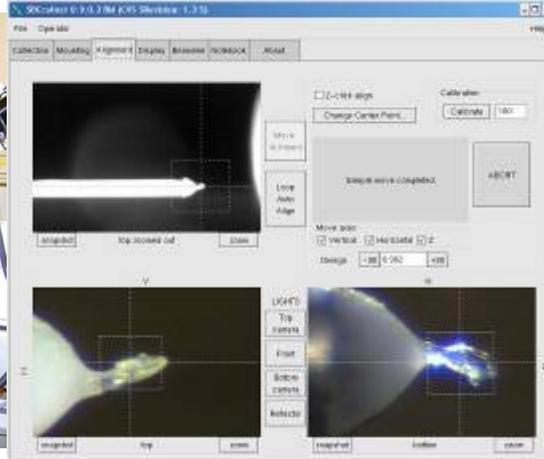
Structural genomics target results: Candidate crossreactors

Target	Score	Rank	Score	Rank	Score	Rank
apc25529	100	1	100	1	100	1
apc25531	100	2	100	2	100	2
apc25534	100	3	100	3	100	3
apc25536	100	4	100	4	100	4
apc25540	100	5	100	5	100	5
apc25544	100	6	100	6	100	6
apc25545	100	7	100	7	100	7
apc25547	100	8	100	8	100	8

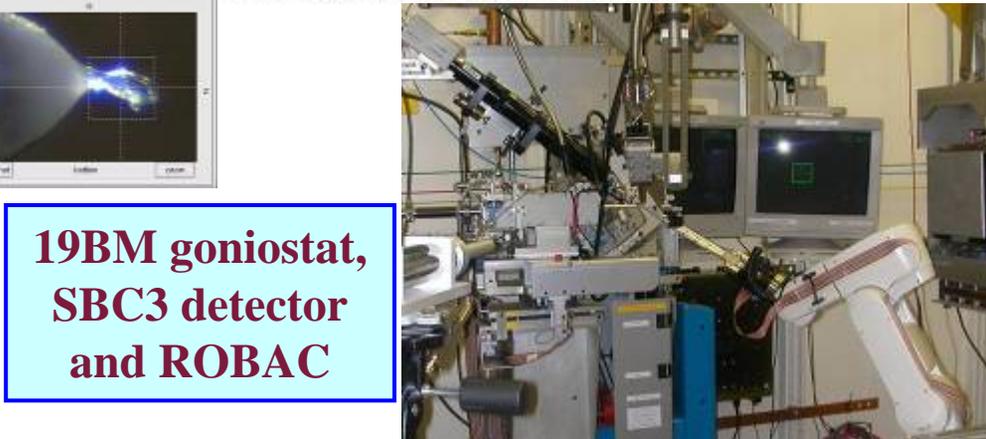
Dedicated X-ray Beamlines for Macromolecular Crystallography - SBC 19ID and 19BM - 1409 PDB Deposits



19ID goniostat and Q315 detector

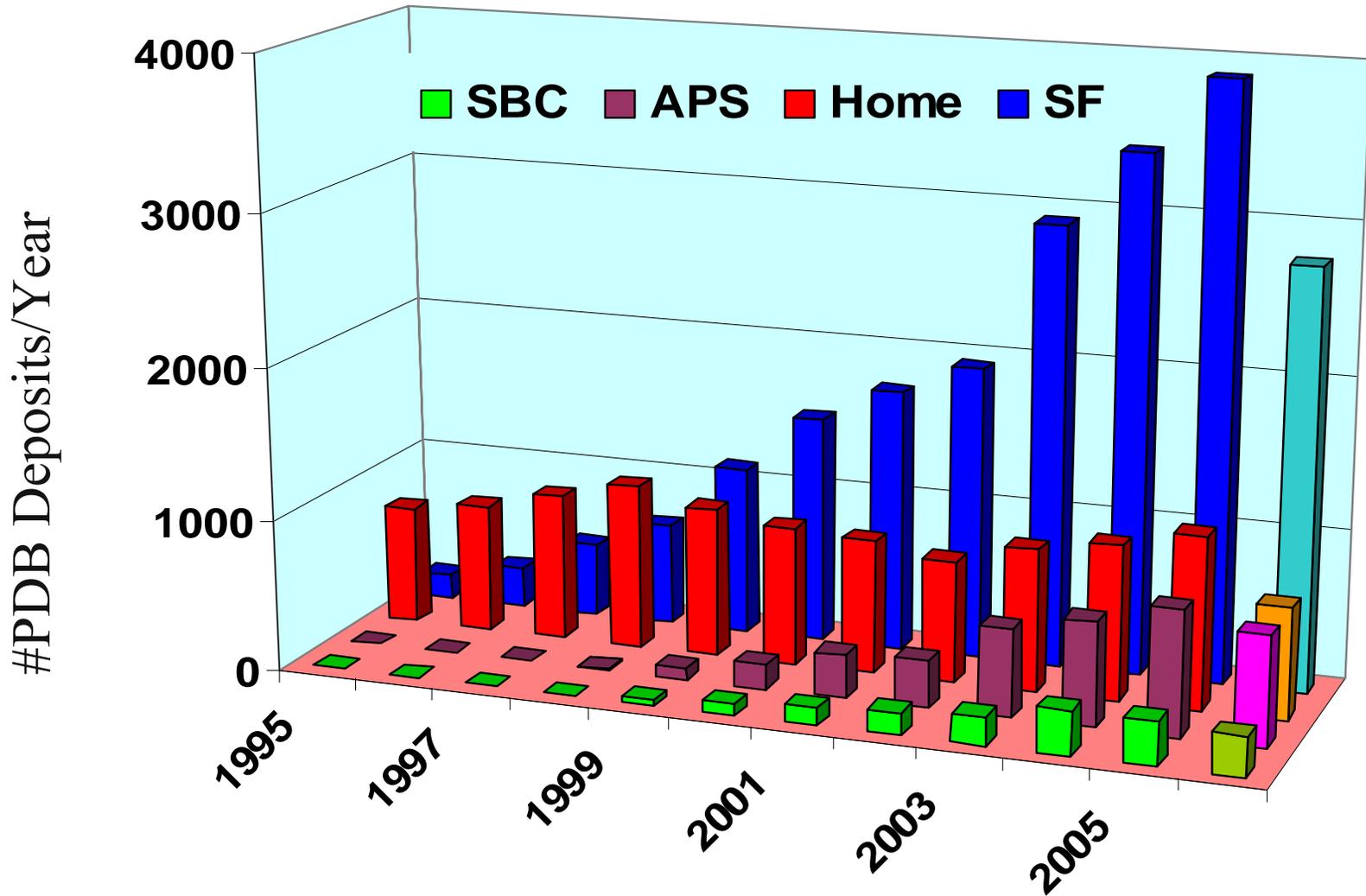


SBCcollect and loop auto-centering



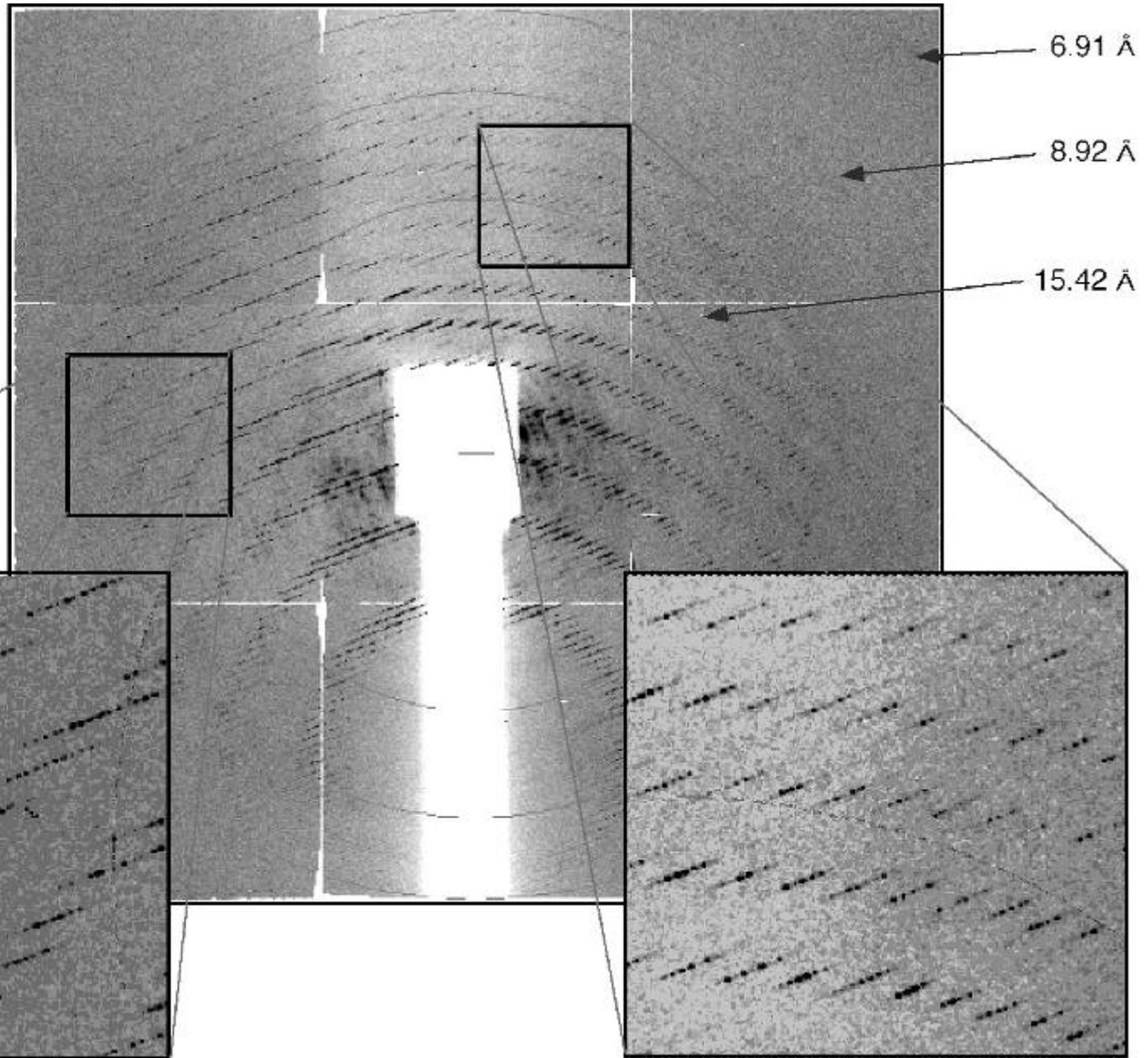
19BM goniostat, SBC3 detector and ROBOT

Impact of Synchrotron Sources and PX Beamlines on X-ray Diffraction Data Collection



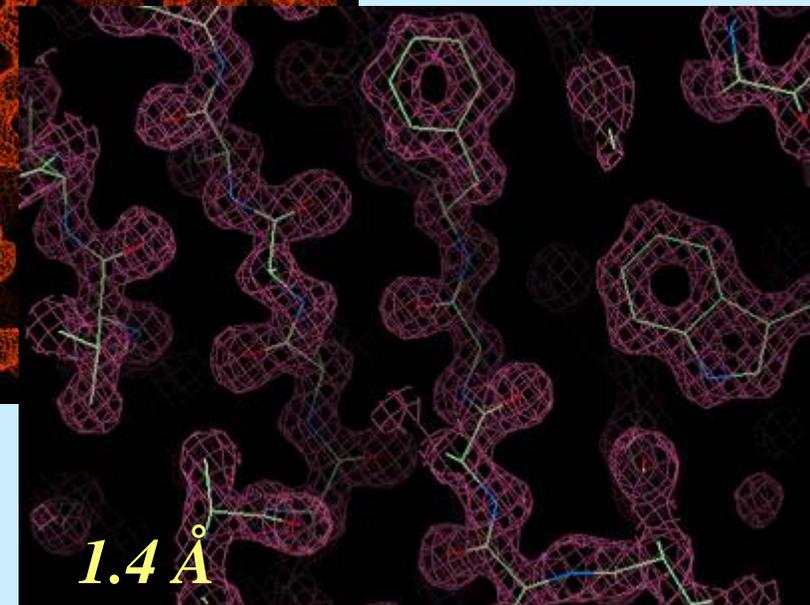
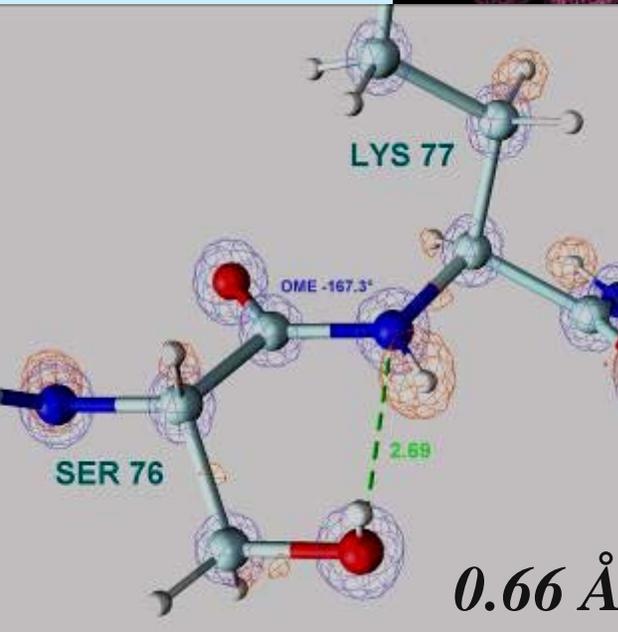
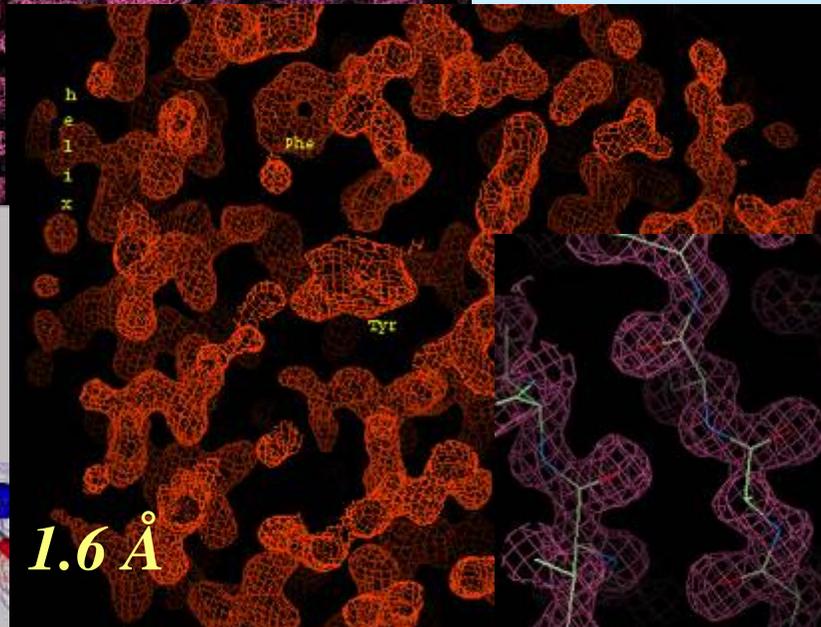
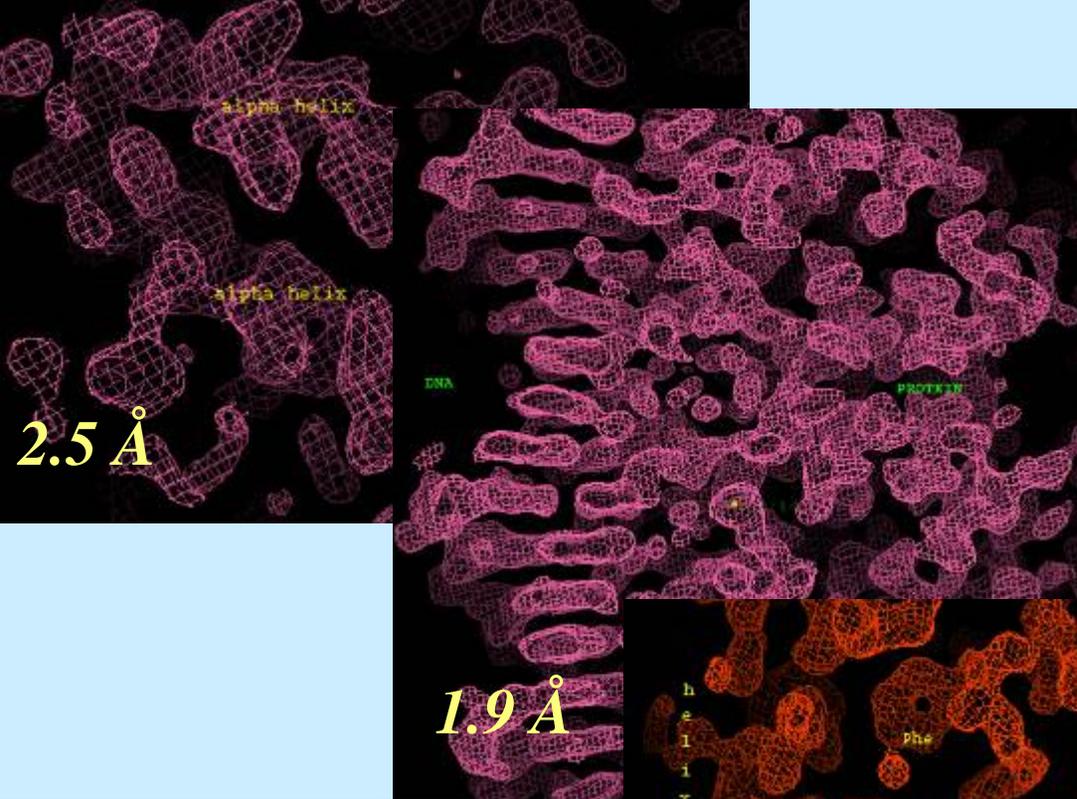
Resolving Very Large Unit Cell

Diffraction pattern obtained at the SBC 19ID beamline from complete ribosomal particles (10 – 12 MDa per asymmetric unit). The crystallographic unit cell is centered, with 2700 Å as the longest axial length. (Courtesy of Dr. Jaime Cate, University of California Berkeley).

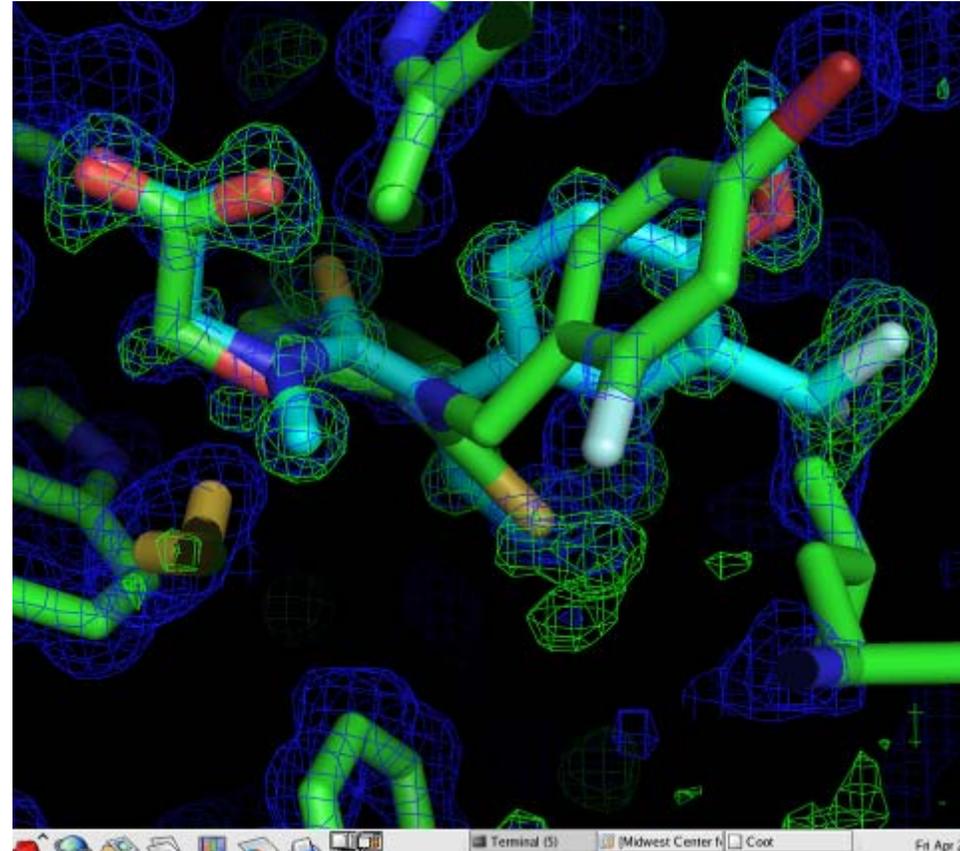
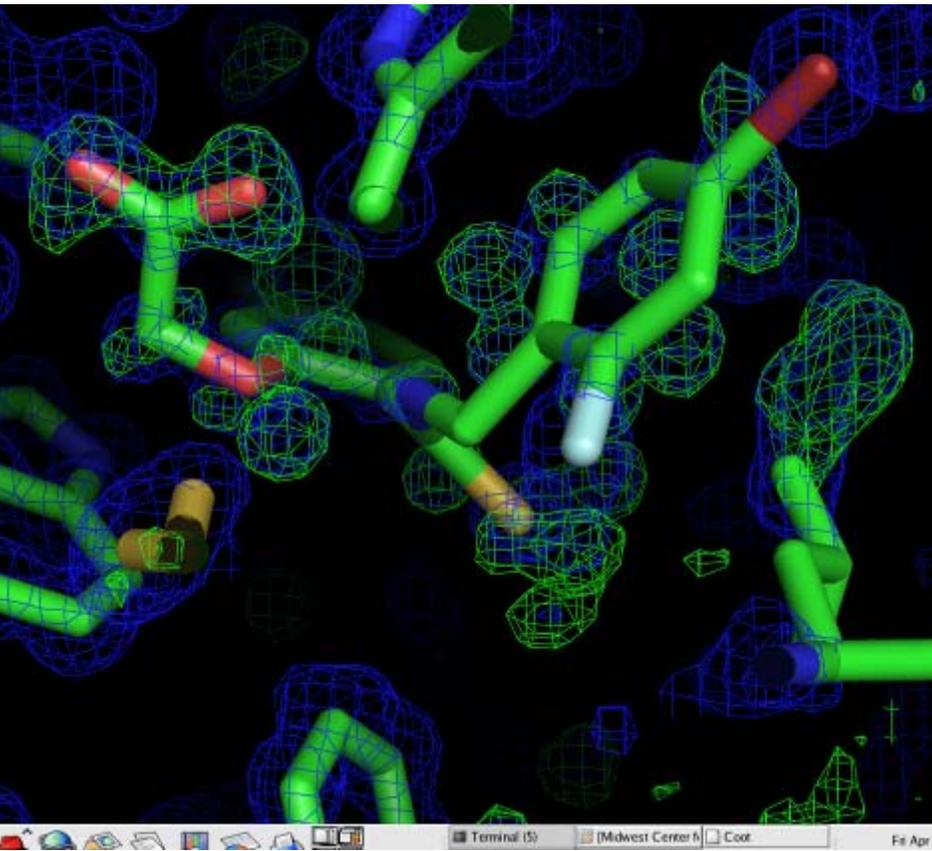


*At 3G SF Crystals
Diffract to the
Resolution Limit*

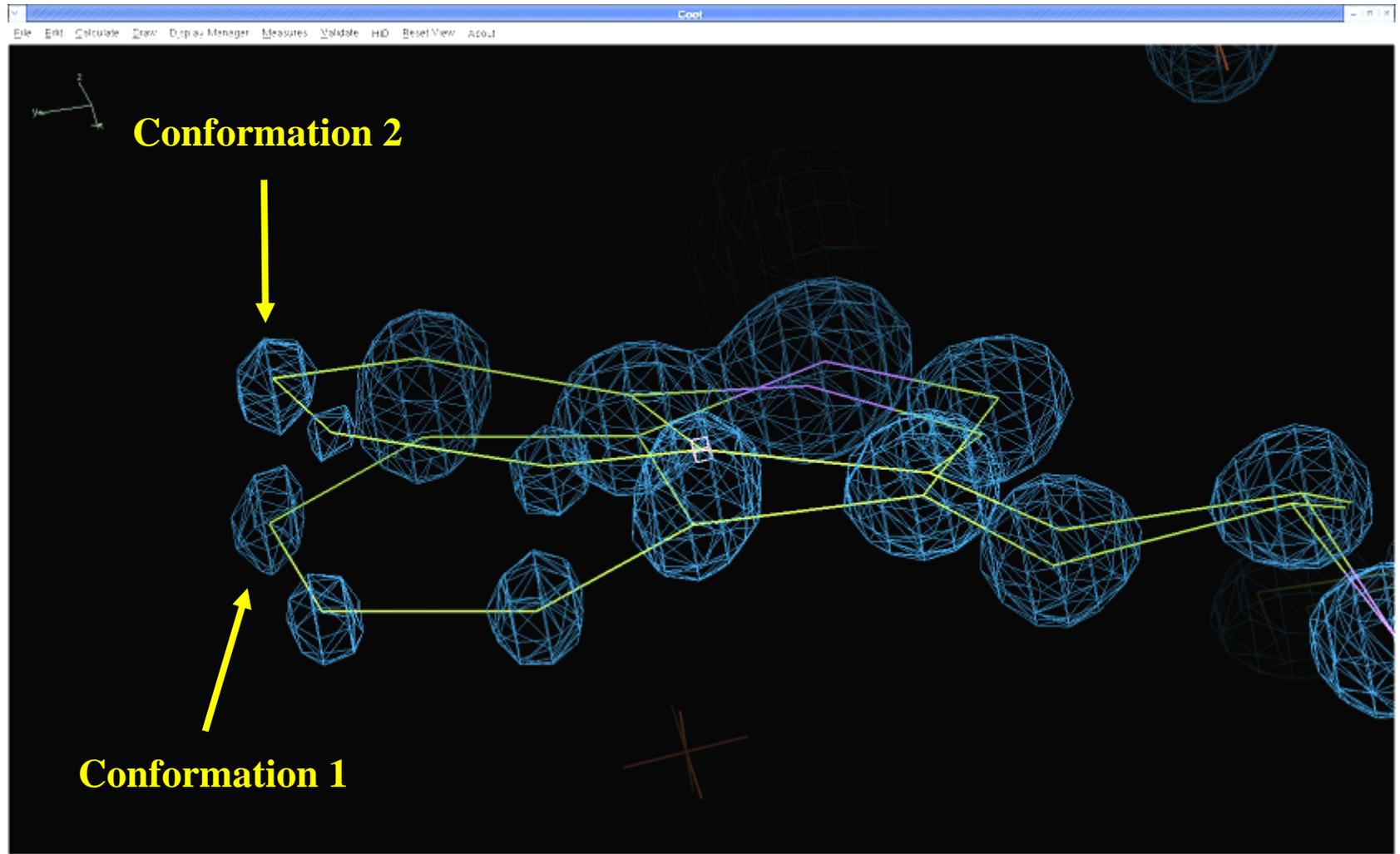
*Electron Density Maps
at Different Resolution
(maps contoured at 1.5σ)*



Details Matter: Binding Two Different Inhibitors in the Active Site of Human Aldose Reductase

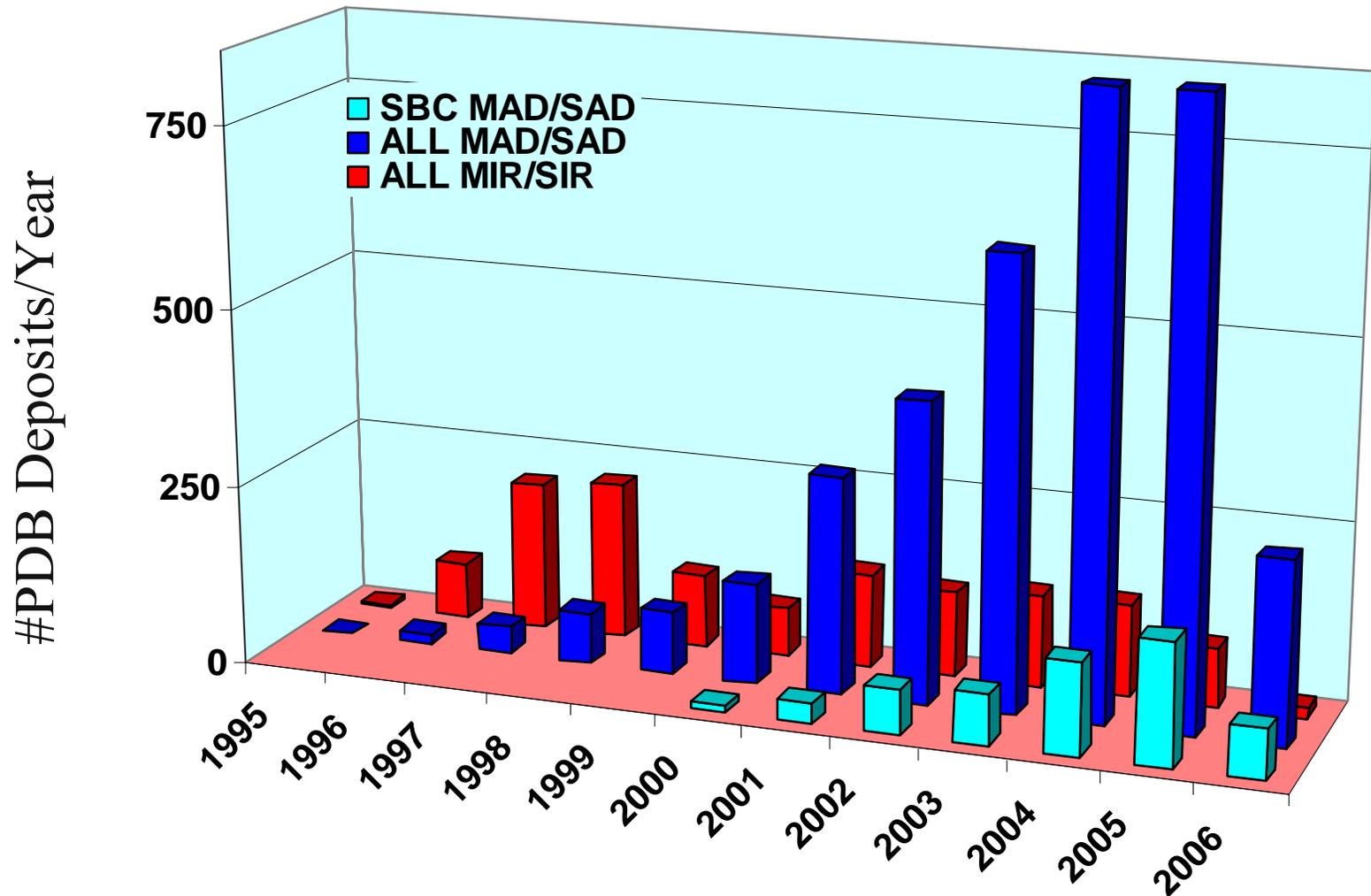


Details Matter: Double Conformation of Trp62 in Lysozyme Structure at 0.65 Å Resolution



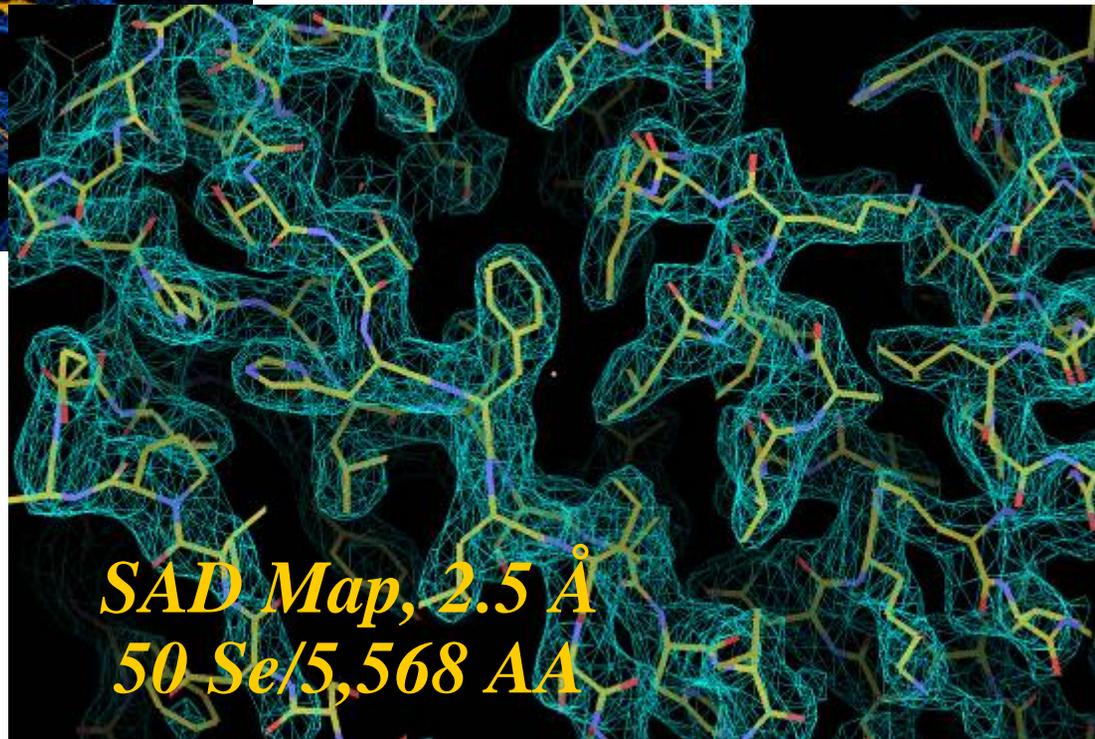
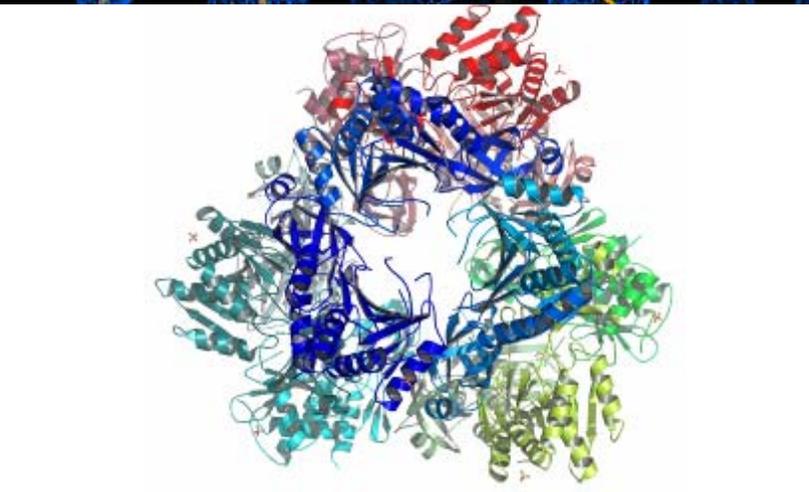
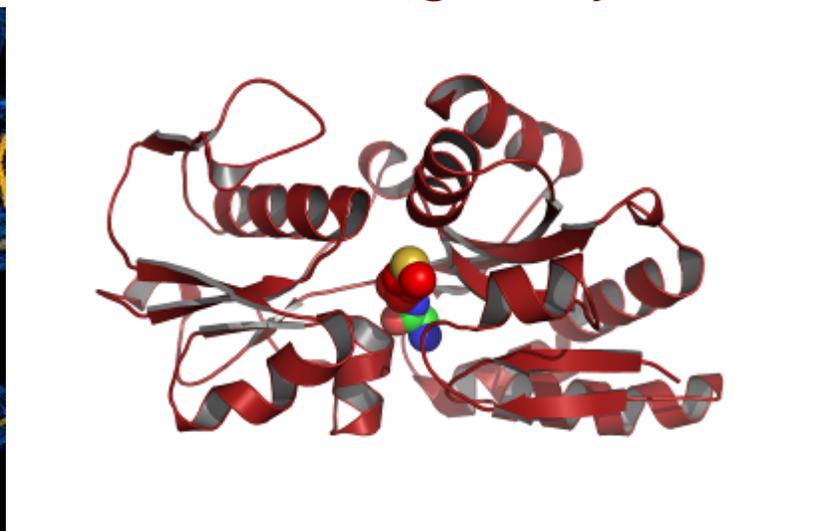
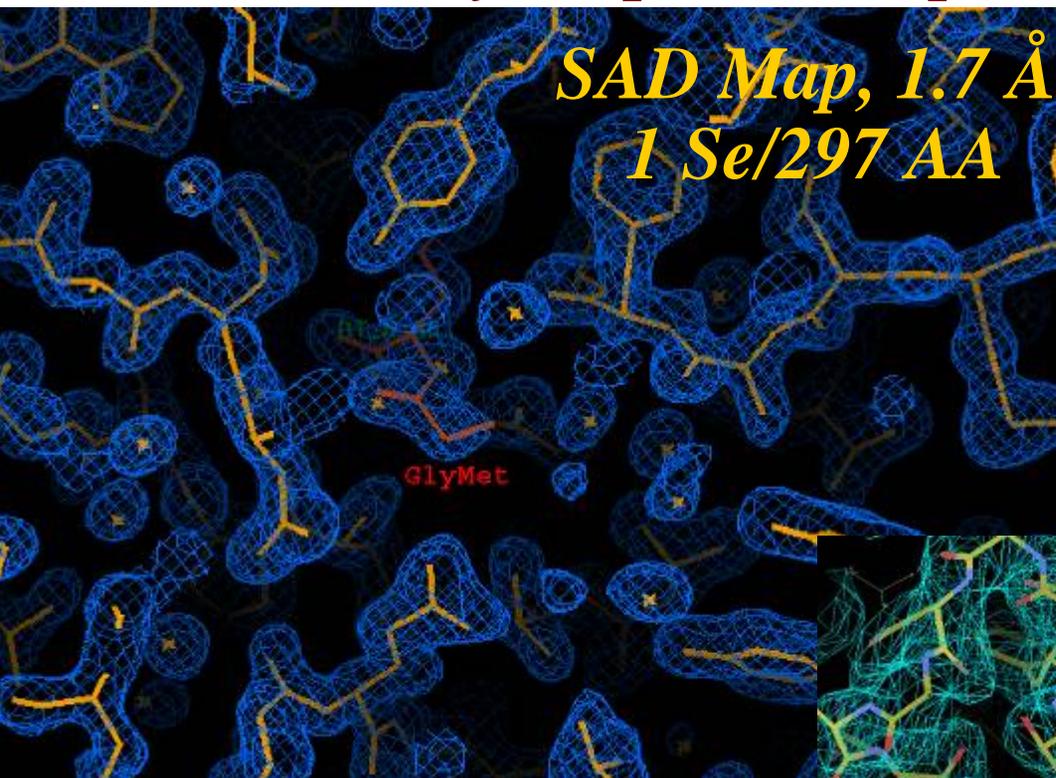
(mol no. 1): C00.6/1*52 TRP occ. 0.42 B 1.74 Hk © per. (C 69,12 91,29 10)

Contribution of SAD/MAD (and SBC) to Determination of New Protein Structures by X-ray Crystallography



46% SBC PDB deposits are solved using MAD/SAD

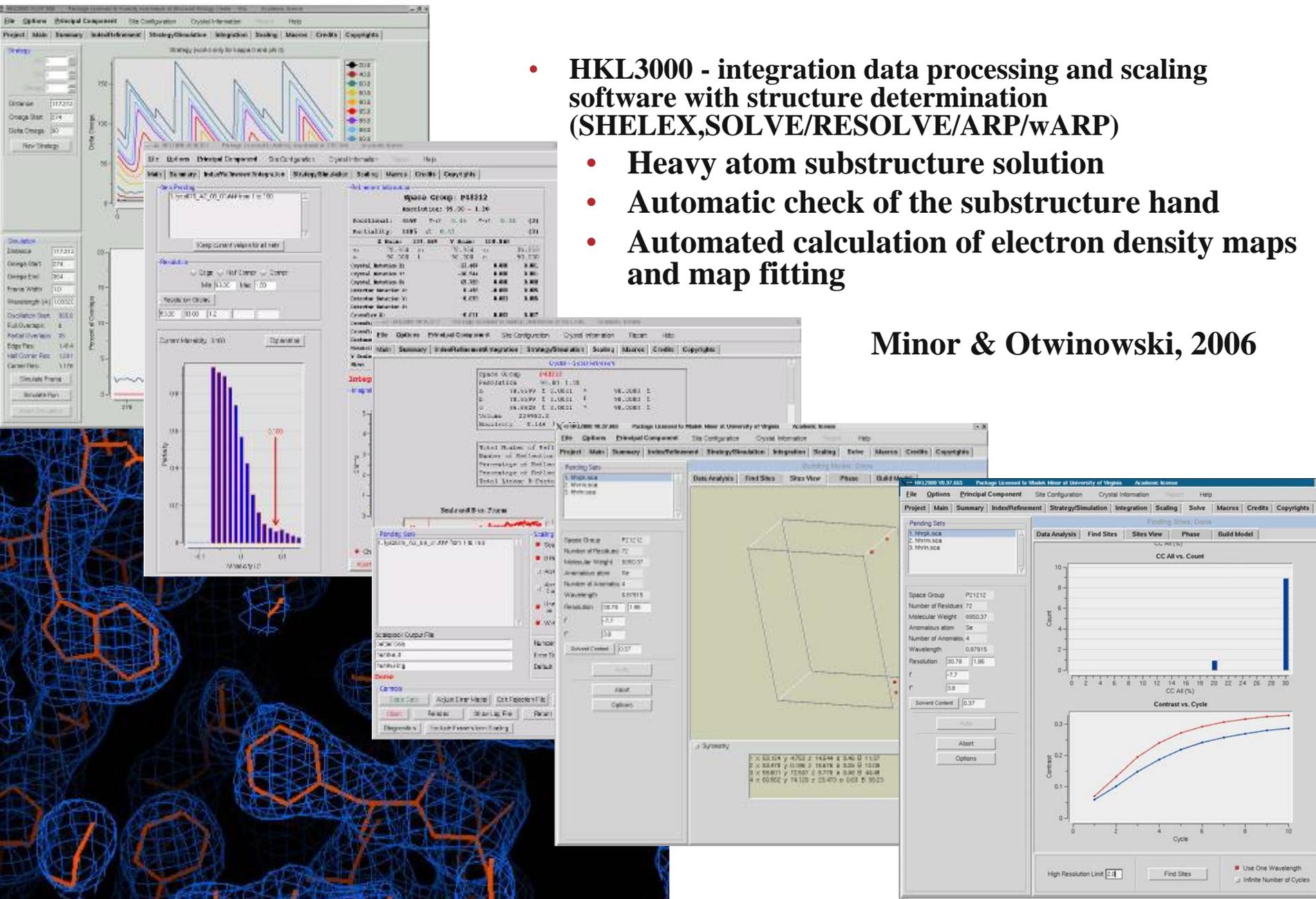
MAD/SAD Phasing Provides Higher Quality Electron Density Maps and Improves Structure Quality



Integration of Data Collection Experiment with HTP *de novo* Structure Determination Tools

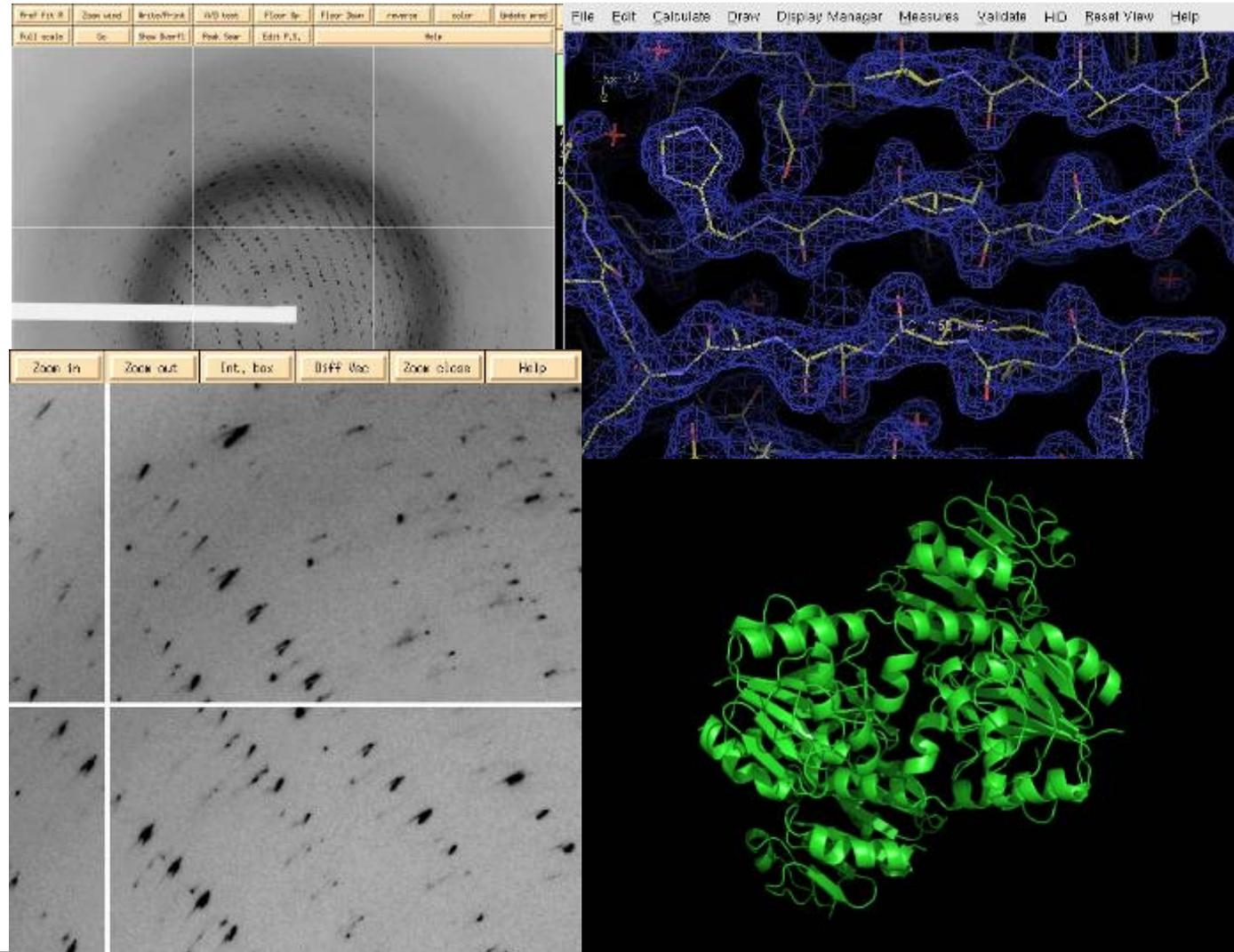
- HKL3000 - integration data processing and scaling software with structure determination (SHELEX, SOLVE/RESOLVE/ARP/wARP)
- Heavy atom substructure solution
- Automatic check of the substructure hand
- Automated calculation of electron density maps and map fitting

Minor & Otwinowski, 2006

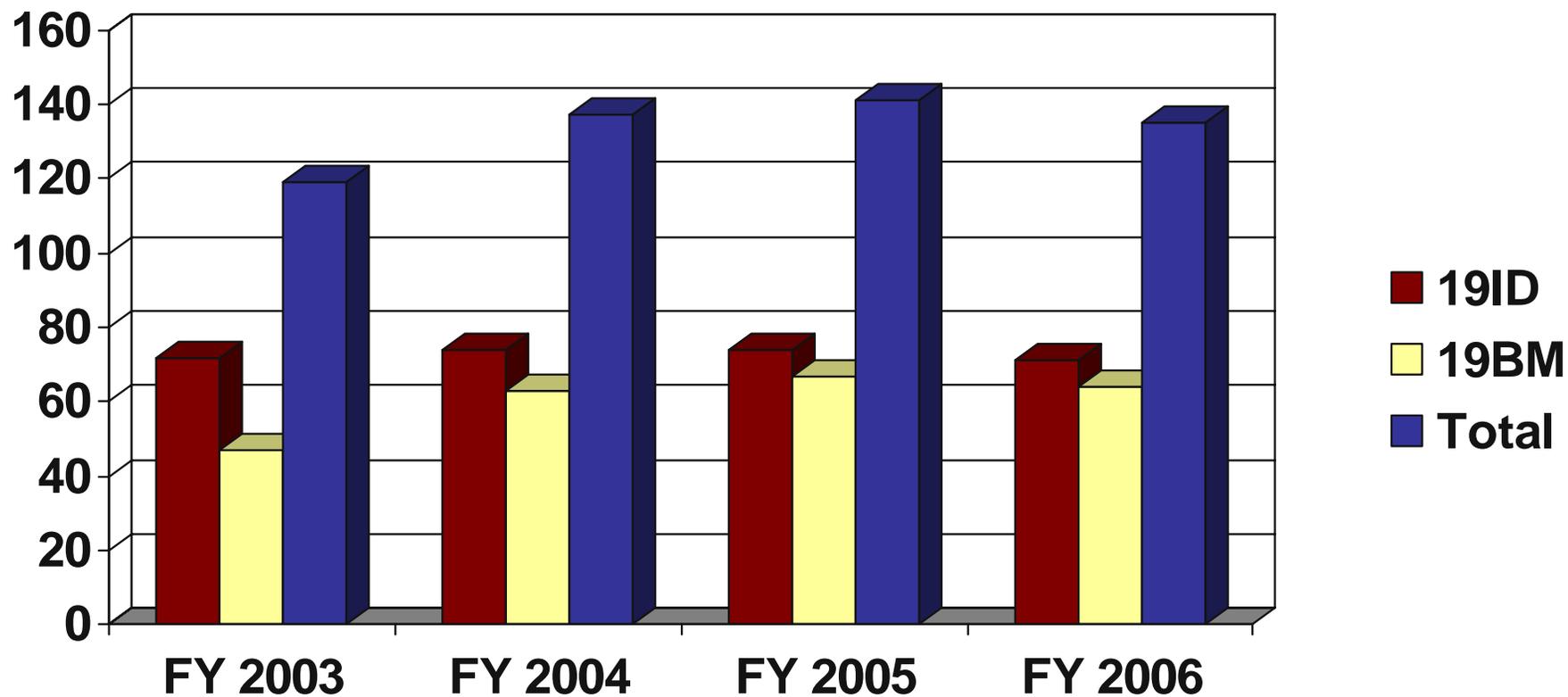


Improving Methods to Handle Difficult Projects - Bad Diffraction

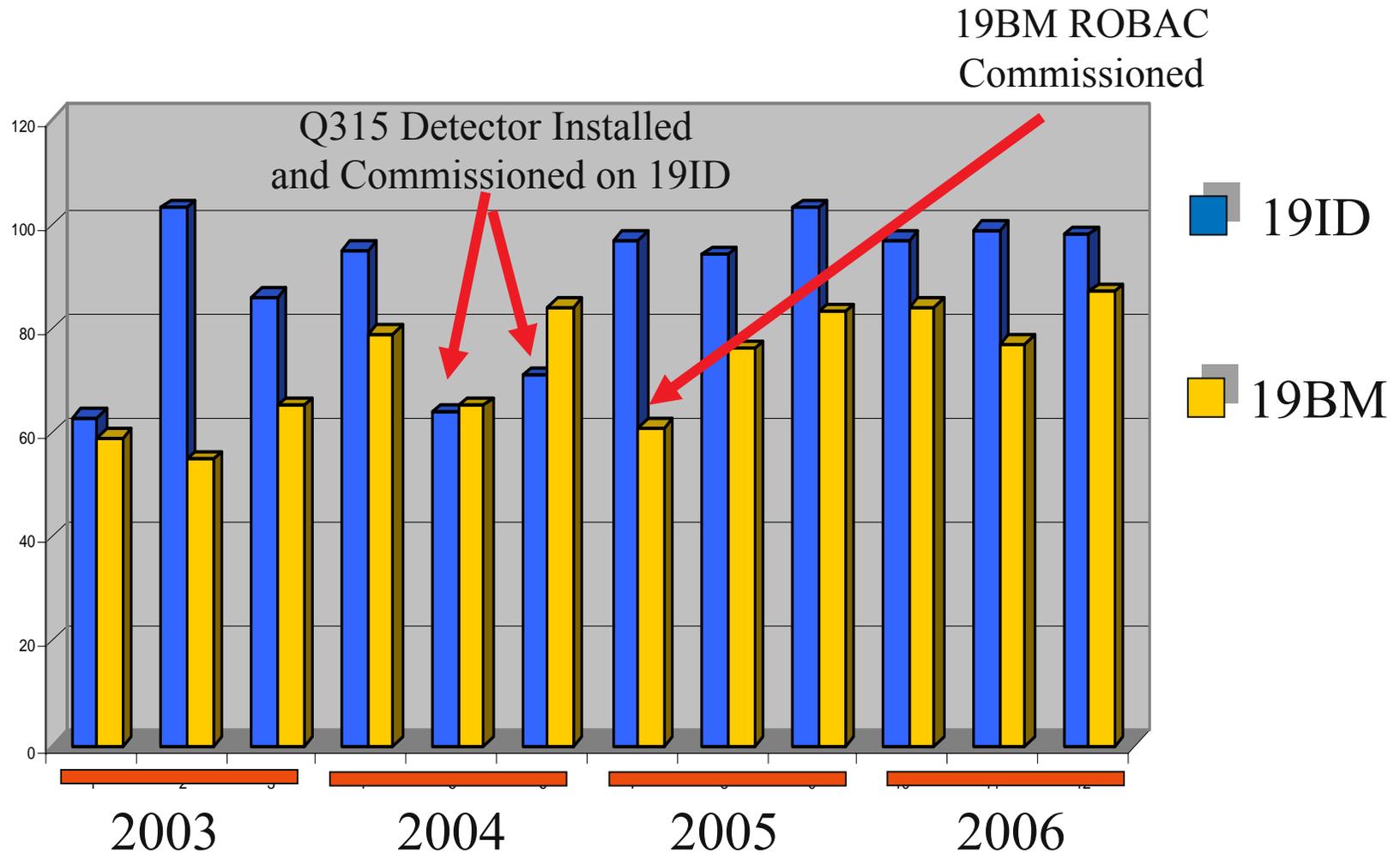
- Space group $P2_1$
- 2.1 Å SAD data were collected at 19ID
- Data processed and solved using HKL2000_PH



SBC User Groups



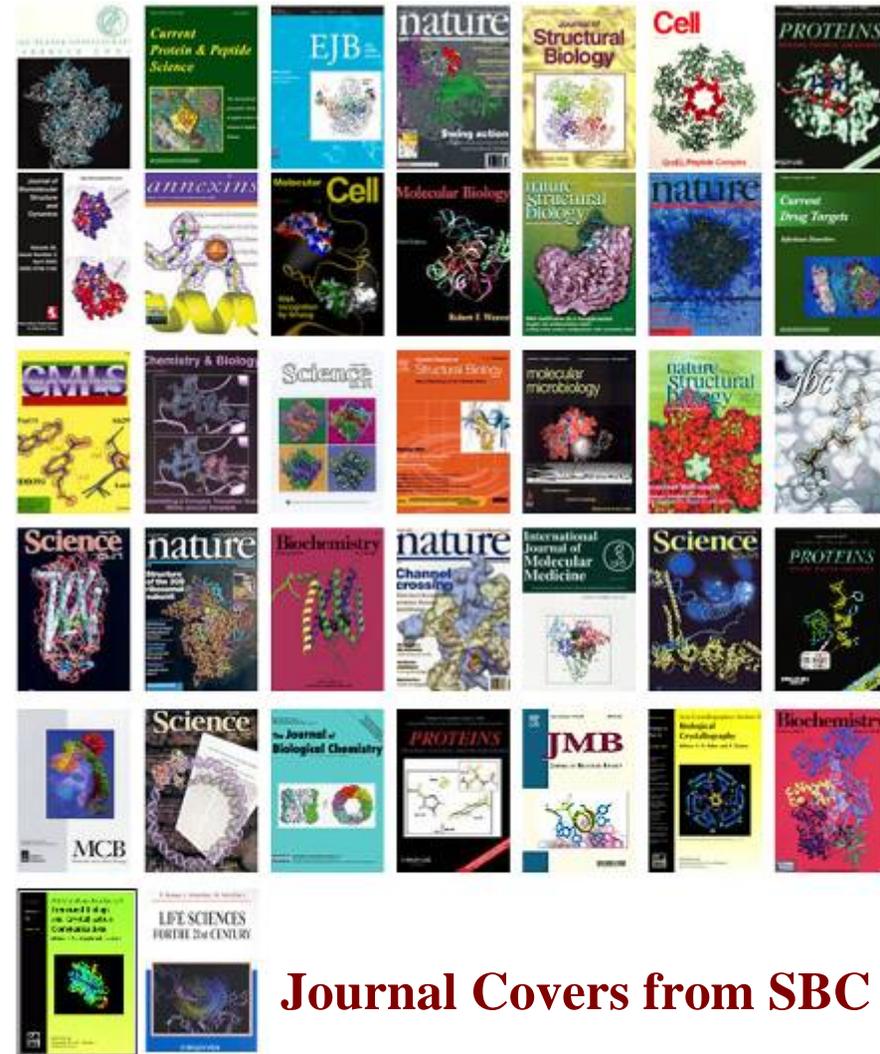
SBC Beamtime Usage



Structural Biology Center Results

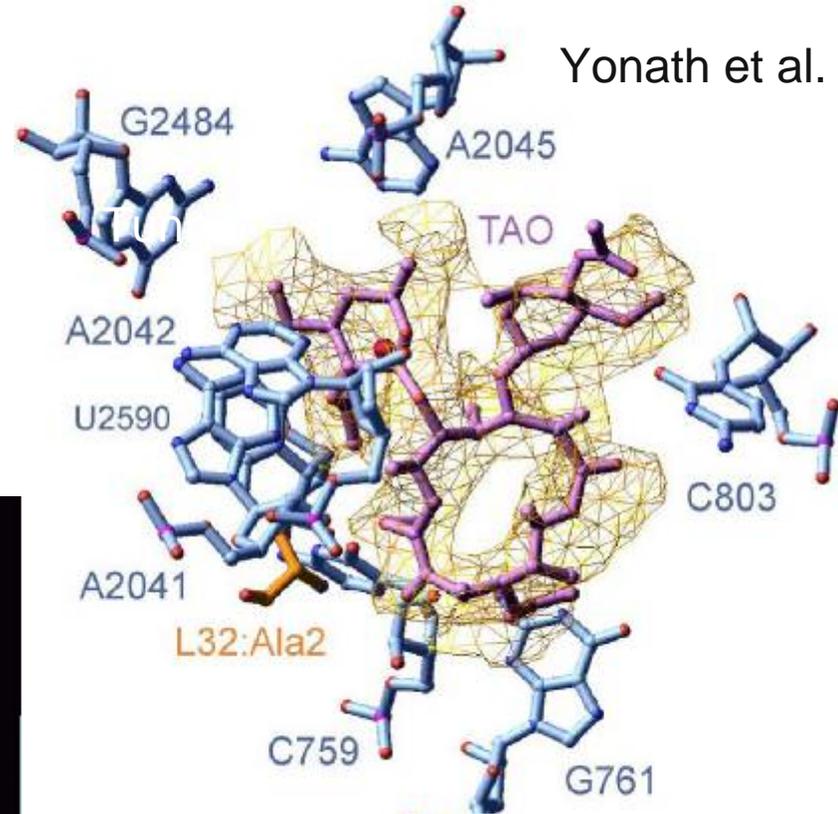
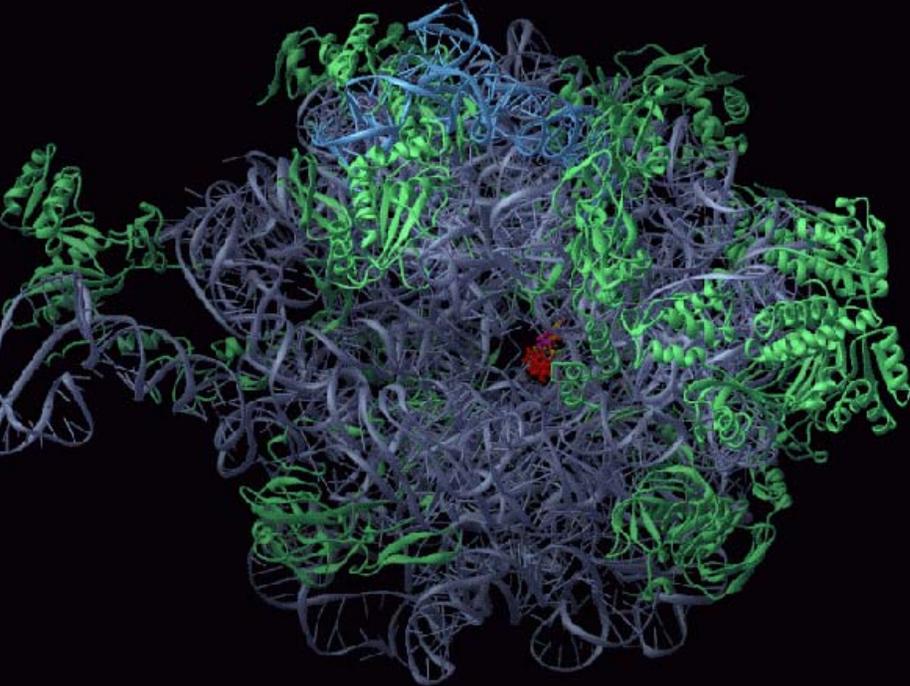
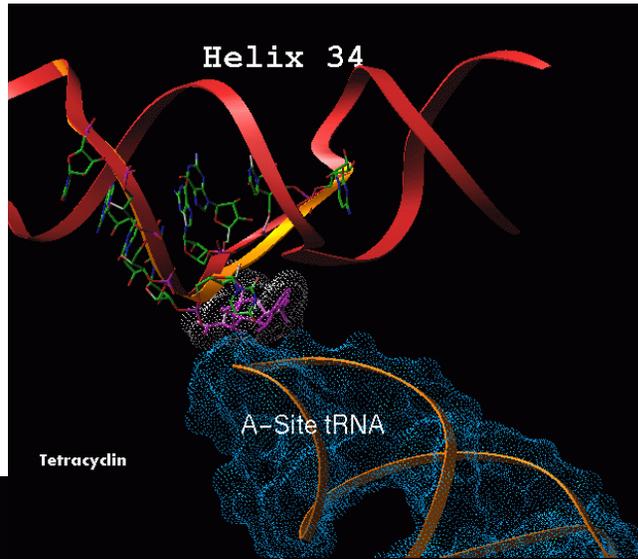
SBC Beamlines Provide a new Approach to Synchrotron Data Collection

- A state-of-the-art facility and user program.
- Efficient use of beamtime is supported by staff, advanced software and hardware including automation in crystal handling, experiment design, data analysis and structure determination.
- SBC staff conducts basic research and develops instruments and methods to address the most challenging problems in macromolecular crystallography.
- SBC staff supports beamline upgrades and construction of new insertion device beamlines with micro-focusing capabilities.
- Data collected at the SBC resulted in 661 publications including 75 papers published in *Nature*, *Science* and *Cell*.

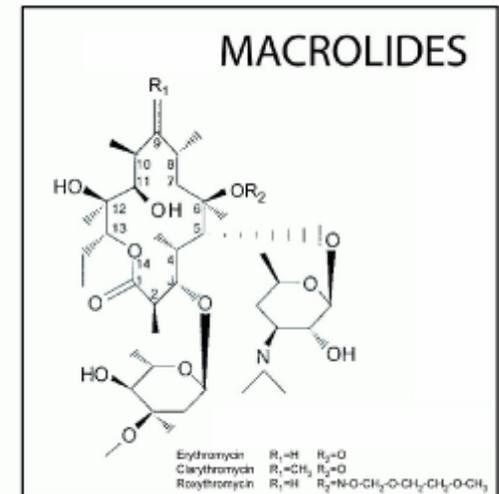


Journal Covers from SBC

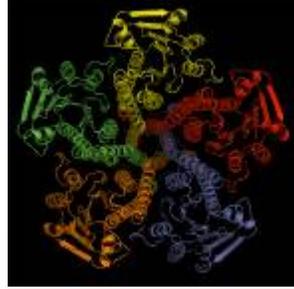
Mapping Binding of Antibiotics to Ribosomal Subunits



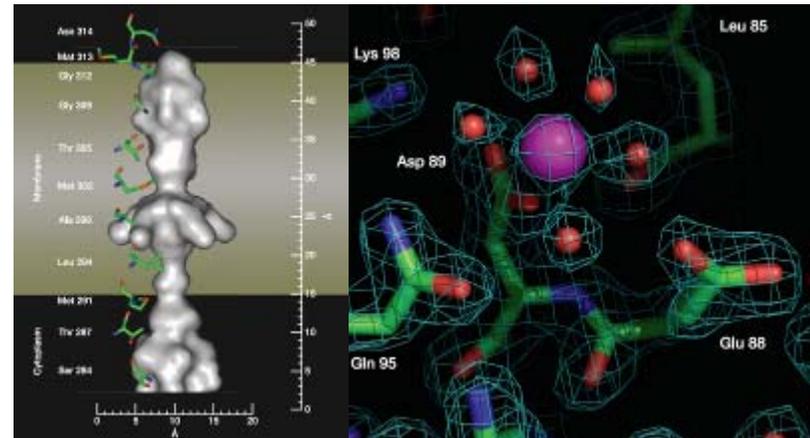
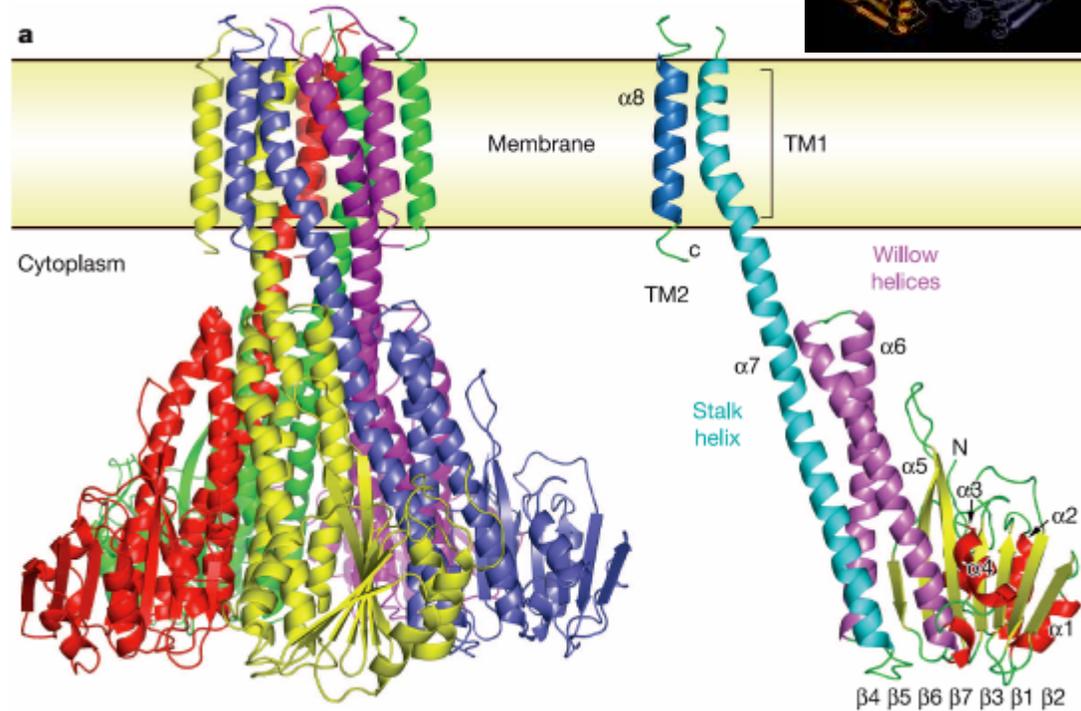
Yonath et al.



Structure of CorA Magnesium Transporter from *Thermatoga maritima*

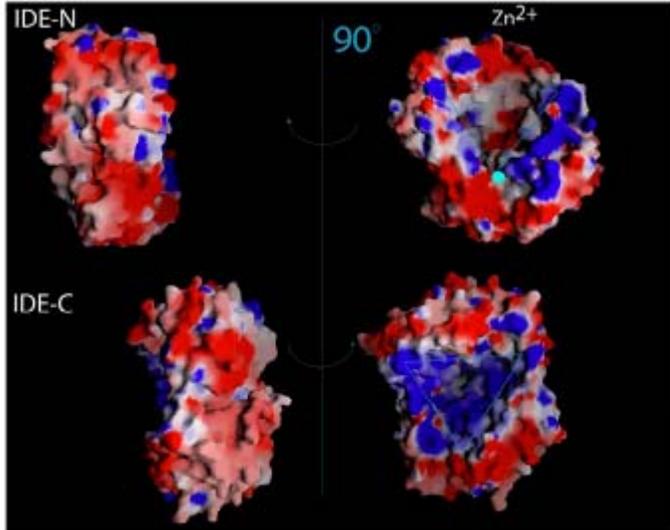
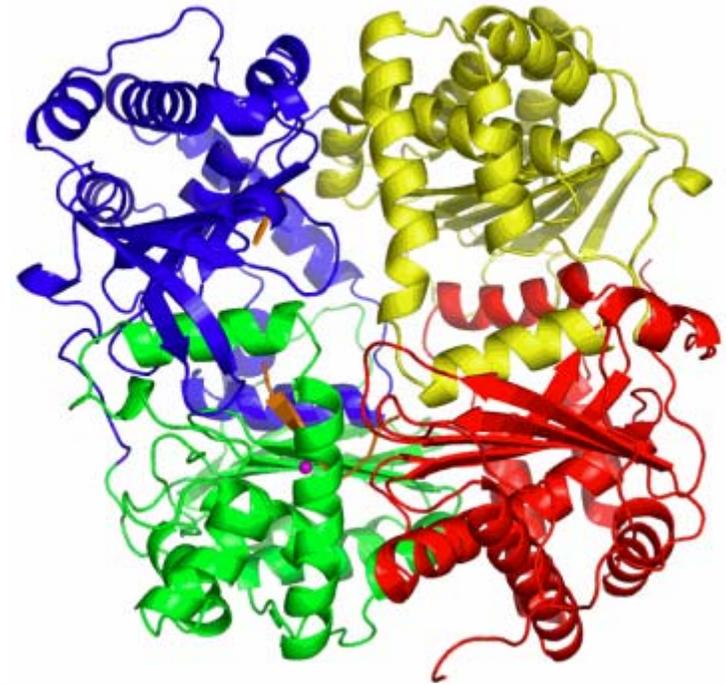
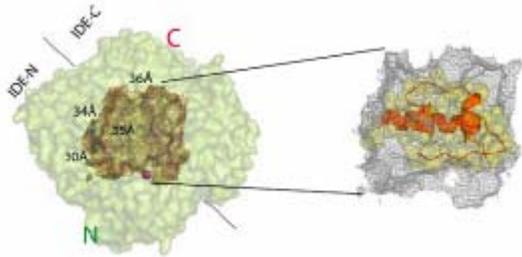


- The CorA magnesium transporters are the primary Mg^{2+} uptake system in prokaryotes and the eukaryotic mitochondrial magnesium transporter. The CorA is widely distributed in the *Eubacteria* and *Archaea*.
- The crystal structures of the entire CorA from *Thermatoga maritima* and of its isolated cytoplasmic domain were determined at 3.9Å and 1.85Å resolution respectively.
- The various CorA homologues have slightly different ion transport specificities. Most CorA mediate Mg^{2+} influx, some Zn^{2+} efflux (Ni^{2+} and Co^{2+} influx leads to toxicity). In eukaryotes, plasma membrane homologues can mediate Mg^{2+} transport and Al^{3+} resistance in yeast and plants.
- The CorA is a funnel-shaped homopentamer with ten transmembrane helices, arranged as on the vertices of two concentric pentagons.



Human Insulin Degrading Enzyme (IDE)

- IDE is evolutionary conserved zinc metalloprotease that is found in bacteria, fungi, plants and animals
- In mammals, 113 kDa IDE preferentially binds and degrades physiologically relevant substrates such as insulin, amyloid β , glucagon, amylin, ANP and TGF- α
- The IDE-knockout mice develop type II diabetes and early symptoms of Alzheimer's disease
- Structures of human IDE with several physiologically relevant substrates have been determined: insulin, amyloid β , amylin and glucagon



Insulin-A chain

GIVEQCCTSI**C**SLYQLENYCN

Insulin-B chain

FVNQHLCGSHL**V**EALYLV**C**GERGFFYTPKT

A β (1-40)

DAEFRHDSGYEVHHQ**KLVFFAED**VGSNKGAIIGLMVGGVV

Amylin

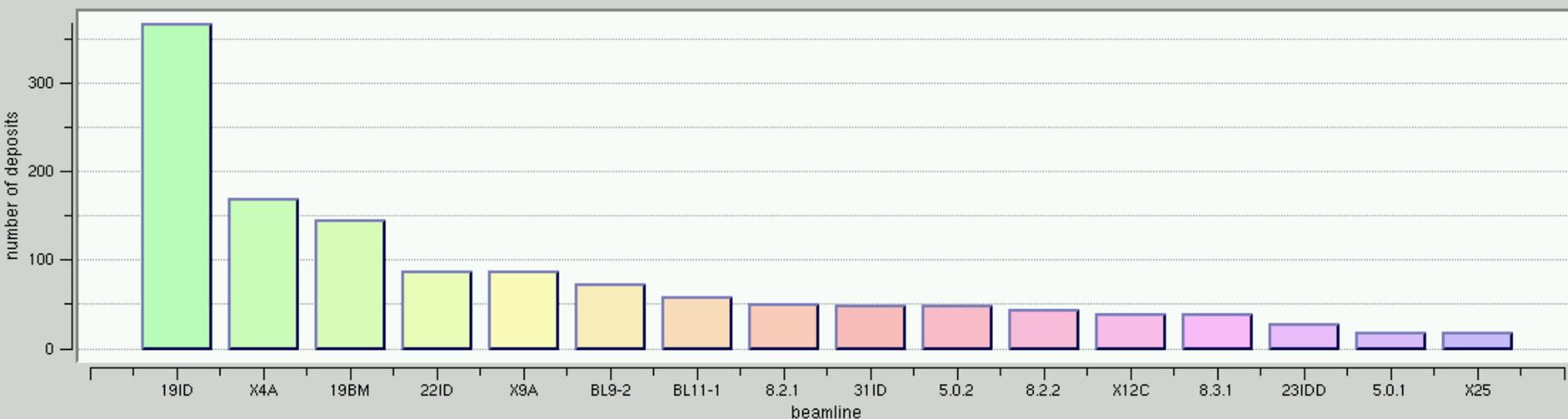
KCNTATCAT**QRLANFLVHSSNN**FGAILSSTNVGSNTY

Glucagon

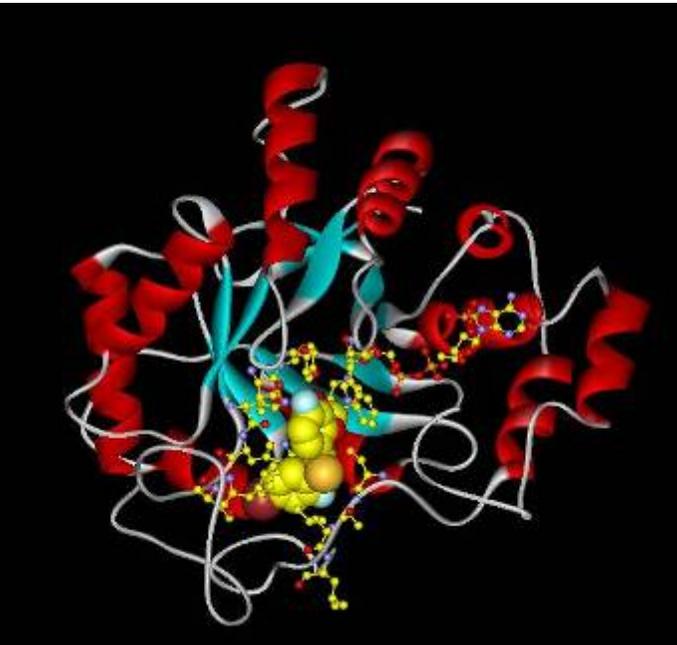
HSQGTFTSDYSKYLD**SRRAQDFVQ**WLMNT

SBC Beamlines also Strongly Contribute to the Protein Structure Initiative

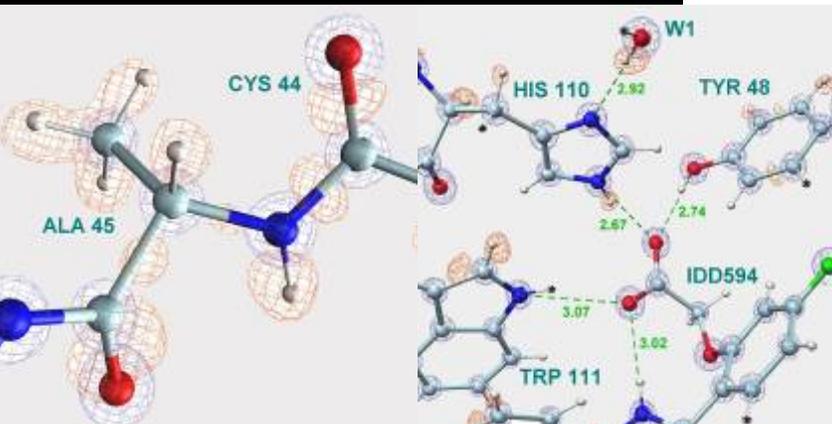
- SBC provided beam time to seven PSI centers
- Interaction of SBC and PSI is very important
 - Crystallographic expertise to support user program
 - Developing and testing high-throughput approaches, automation and robotics
 - No shortage of crystals and projects for beamline testing
 - Better use of beam time
 - Nearly instant access for testing crystals



Aldose Reductase - Data Collection at Atomic Resolution



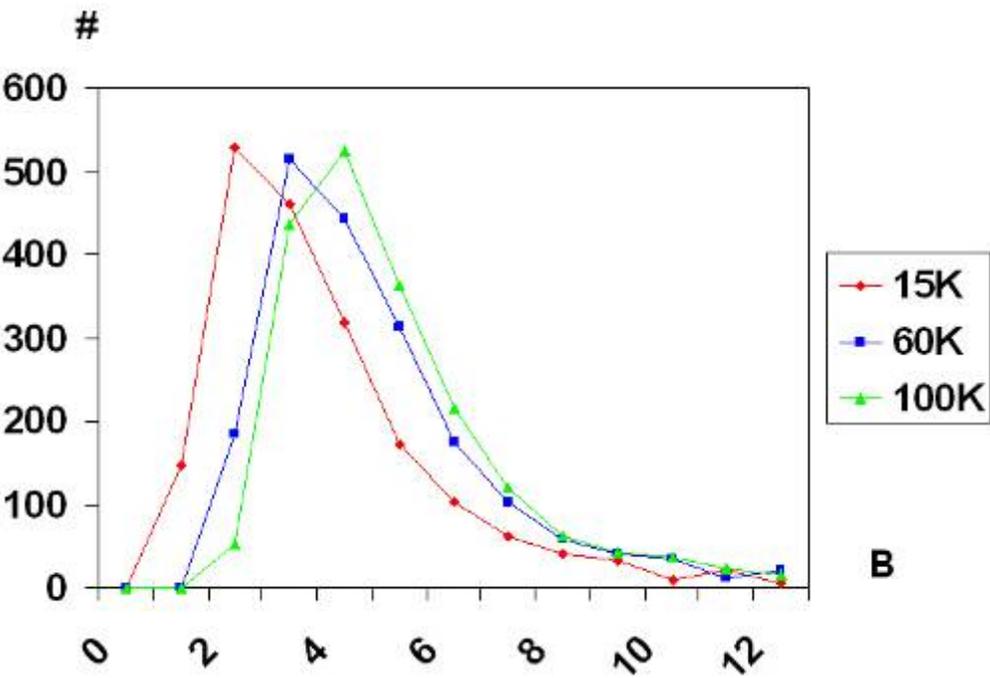
- Aldose reductase is an excellent model system
 - NADPH-dependent enzyme that reduces a wide range of substrates, such as aldehydes, aldoses and corticosteroids
 - Medically relevant - reduction of D-glucose into D-sorbitol is believed to cause severe degenerative complications of diabetes
 - 315 amino acid residues, $(\alpha/\beta)_8$ TIM barrel fold
 - The ternary complex of human aldose reductase-NADP⁺-inhibitor is of significant therapeutic interest
 - Crystals of human aldose reductase complexed with inhibitors and NADP⁺ diffract to 0.6 Å



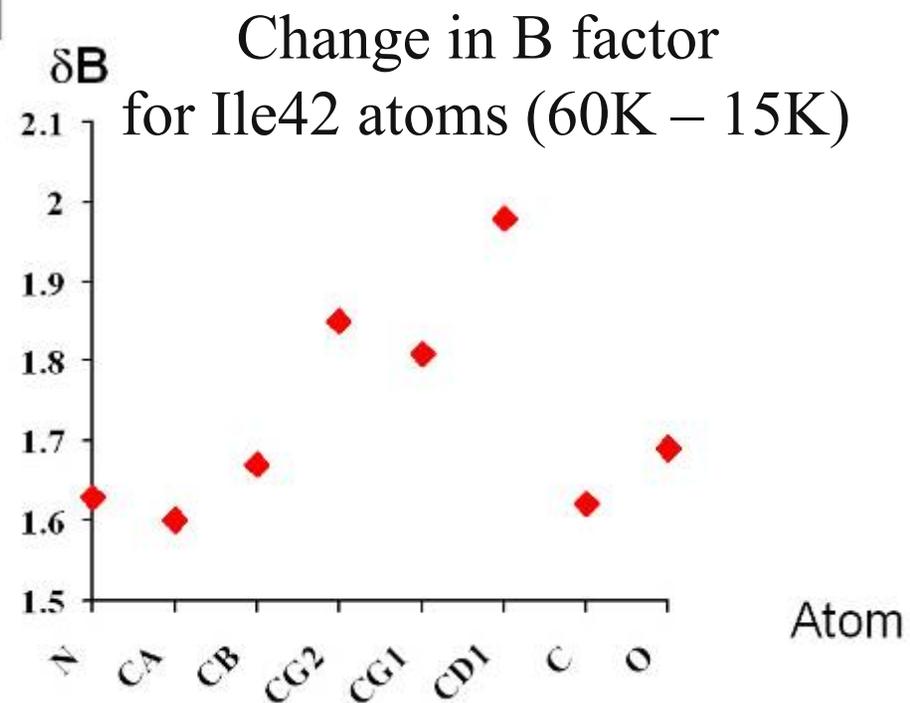
Results:

- Structure of aldose reductase determined and refined at 0.66 Å
- MAD phases obtained at 0.86 Å resolution
- Experiments at helium temperatures – studies of how temperature contribute to crystallographic B factors
- Radiation damage experiments

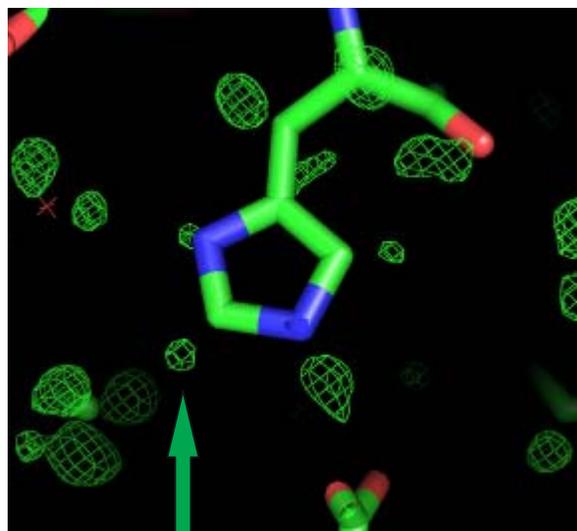
Structure Determined at Liquid Helium Temperature Shows Lower B-factors for Many Atoms



Change in B factor for aldose reductase atoms

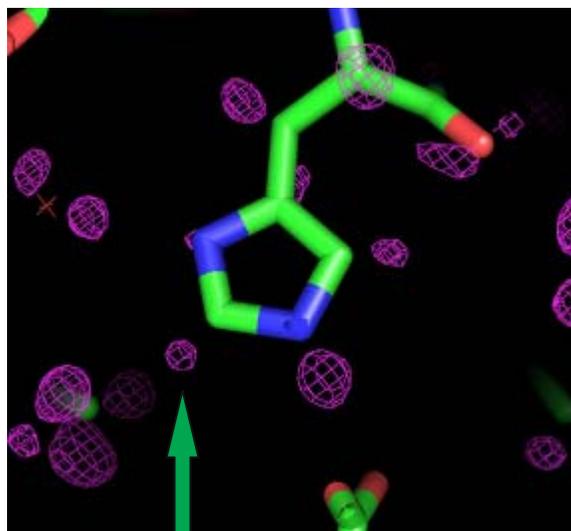


Experiments at Liquid Helium Temperatures: Hydrogen Atoms of His110 in the Active Site of Aldose Reductase



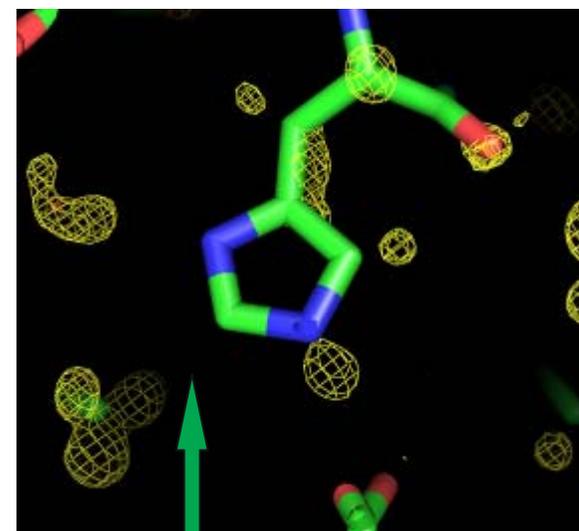
15K

Hydrogen atom



60K

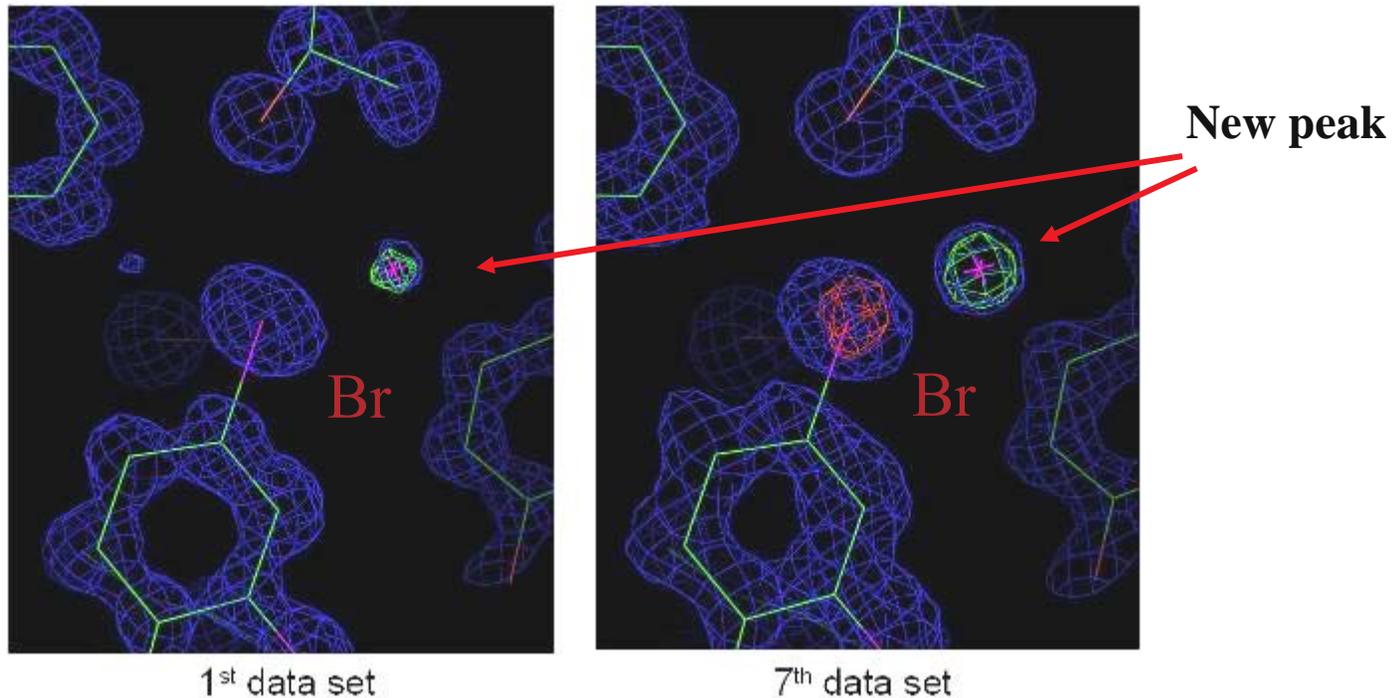
Hydrogen atom



100K

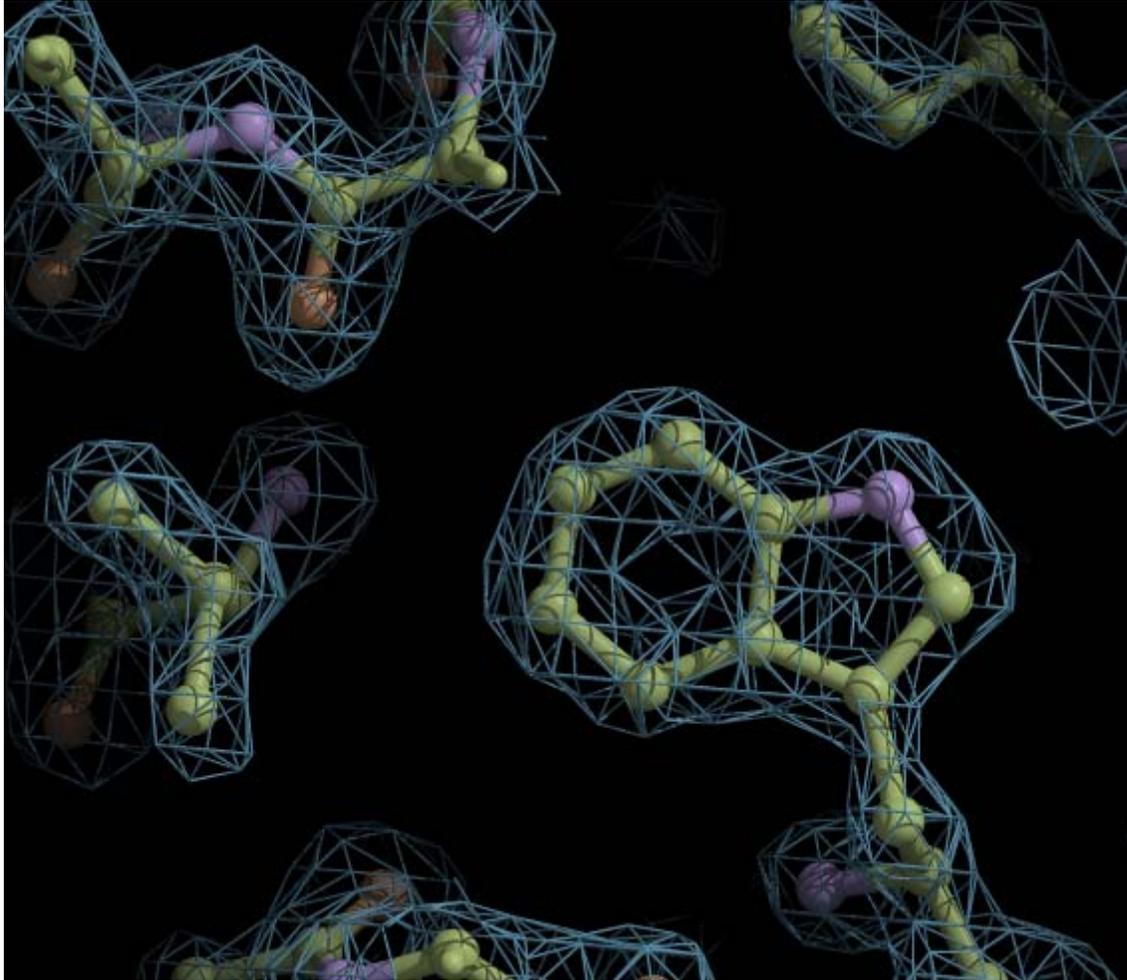
Missing hydrogen atom

Radiation Damage - Studies of Aldose Reductase at Atomic Resolution



Radiation damage induced during data collection in the crystal of human aldose reductase. The electron density of Br atom diminishes with x-ray dose and new peak appears nearby and increases with x-ray dose. Red peak at the second picture indicates that occupancy of Br atom has decreased, green peak nearby has increased. Estimated occupancy of Br = 0.8 (1st data set) and 0.44 (7th data set). (blue contour - 1σ , green contour - $+8\sigma$, red contour - -8σ level of Fobs-Fcalc map).

1.9 Å Structure of Lysozyme Solved at 19BM Beamline Using Sulfur (and Chlorine) Anomalous Scattering



Space group: $P4_32_12$

Data collected at 19BM, E=7 keV, (1.7694 Å)

Resolution: 1.9 Å

HKL3000 found 16 sites of S and Cl,

ARP/wARP autotraced 128/129 residues

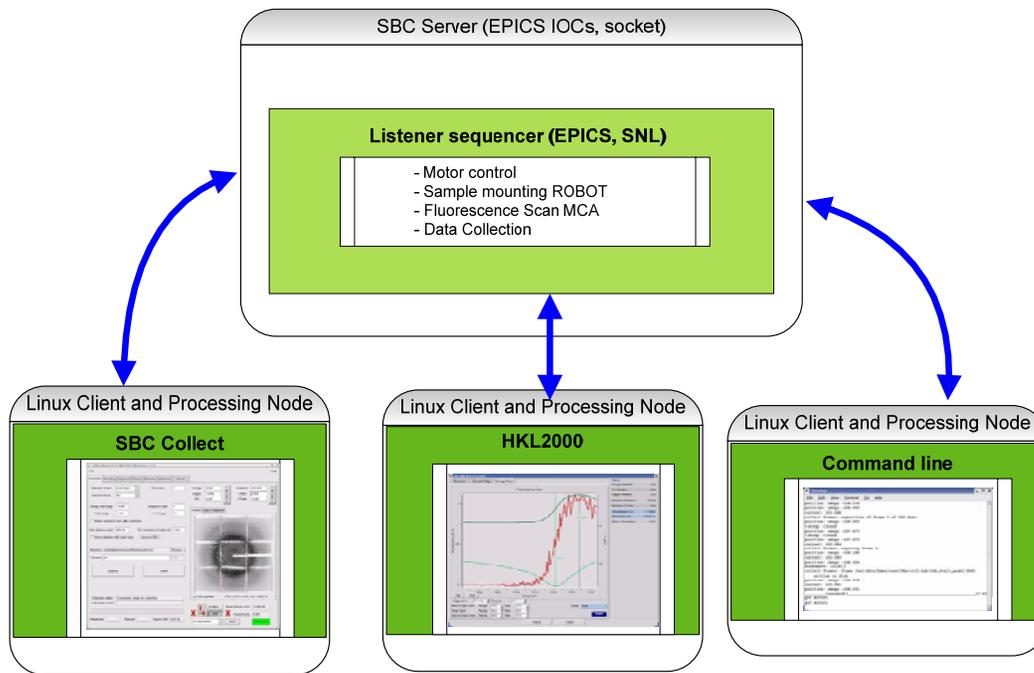
R-factor/Rfree: 16.9%/21.3%

Integration of SBC Software - SBCserver

- Runs through inetd/xinetd, modular architecture
- C, C++, ca (native), pvca, seq (SNL)
- Asynchronous communication
- Security through login and password
- Multi-platform

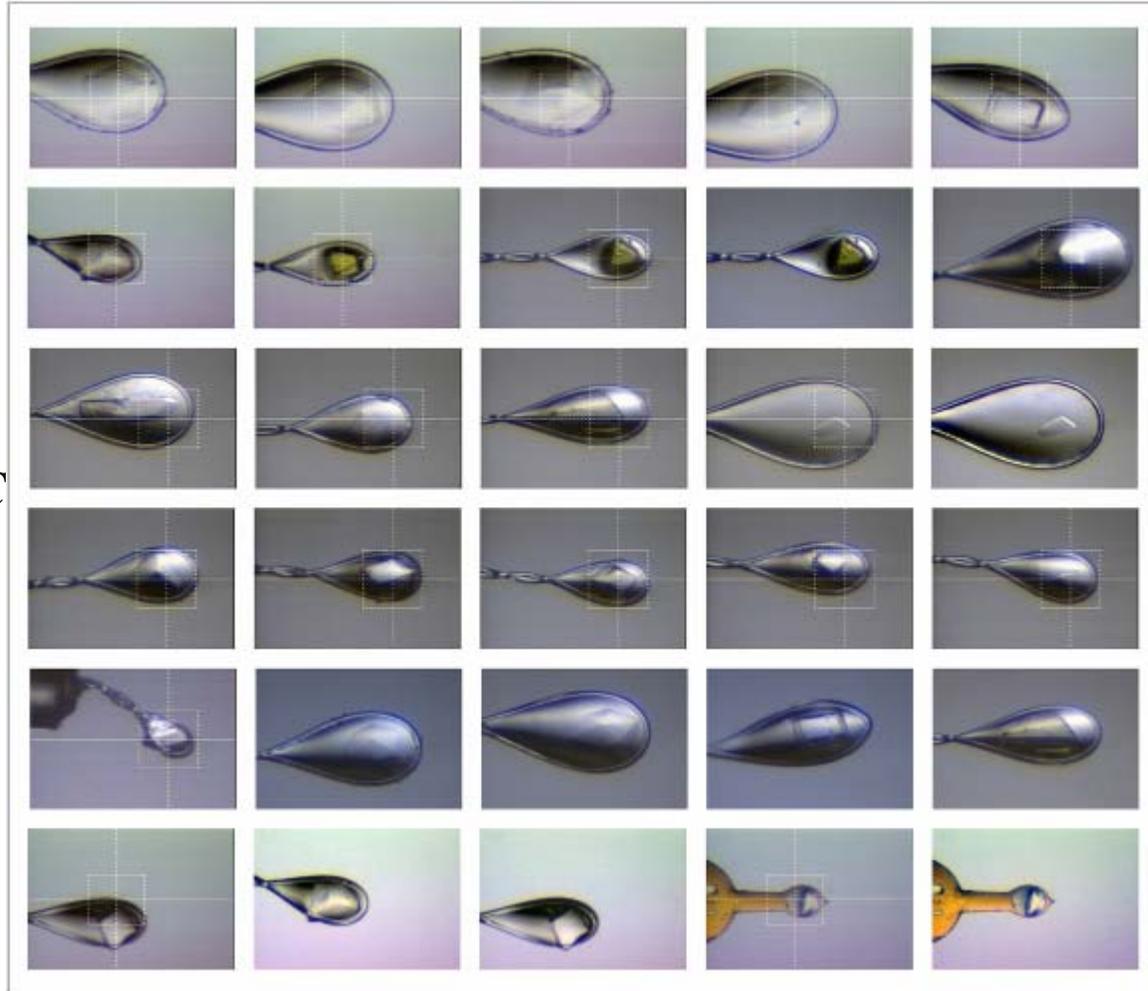
- **The SBCserver controls or monitors the following objects:**

- Goniostat motion system
- Detector support
- Beamstop motion system — x, y, z movements
- Slits motion system
- Timing shutter
- Crystal alignment motion system — x, y, z movements
- Energy (wavelength) of the X-ray beam
- Fluorescence spectrum recording
 - *using EPICS scan record*
 - *automatic attenuation*
- Beam intensity monitoring (ionization chambers)
- Hutch door status
- Beam status
- Crystal-mounting robot
- PSS (personal safety) system

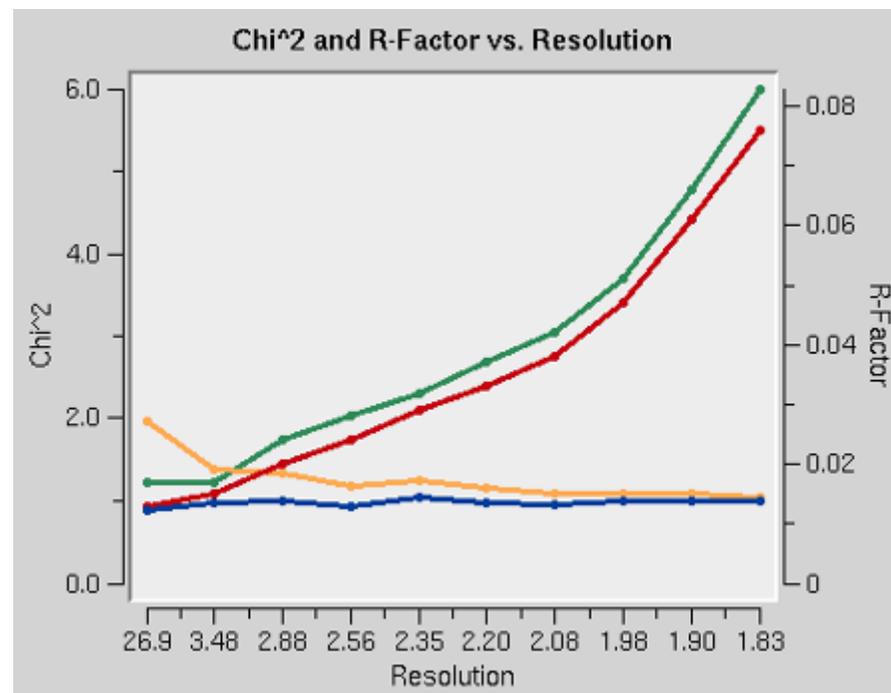
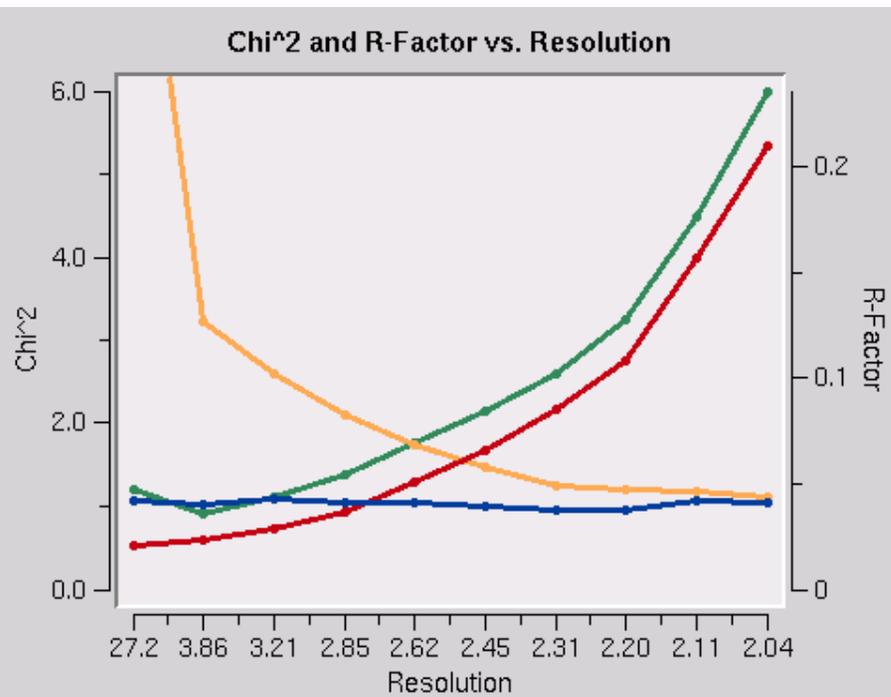


Data Acquisition at SBC Beamlines

- Data acquisition
 - Crystal mounting
 - *Robot*
 - **Crystal alignment**
 - *Loop centering - SBC in house software*
 - *Crystal centering XREC software EMBL (V. Lamzin)*
 - Crystal pre-screening
 - Evaluation for further data collection
 - Data Collection



HKL3000 - Anomalous Signal Detection



Integration with Databases Quick Check Against PDB

The screenshot shows a software interface with several windows. The main window has tabs for 'Project', 'Main', 'Summary', 'Index/Refinement', 'Strategy/Simulation', and 'Integration'. The 'Summary' tab is active, showing 'Protein Data' for a protein with Swiss-Prot accession code AAB84656, name AAB84656, and description 'conserved protein [Methanother...]'.

A 'Summary' window is open, displaying search results for PDB entries with high identity:

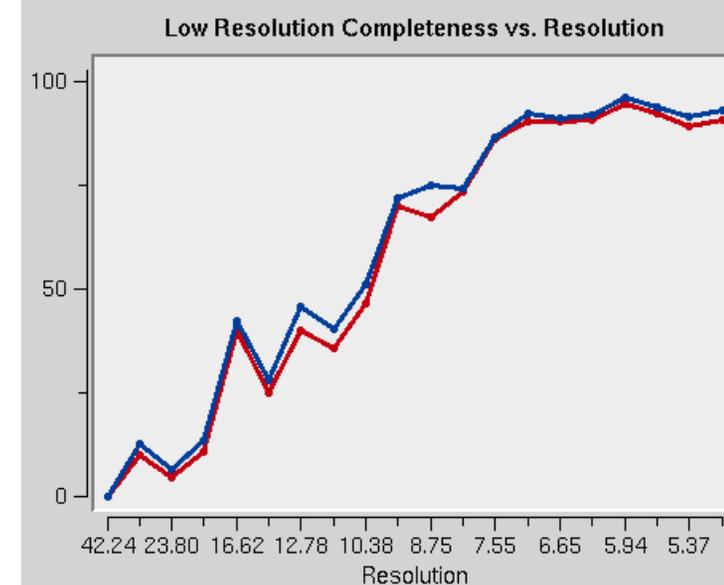
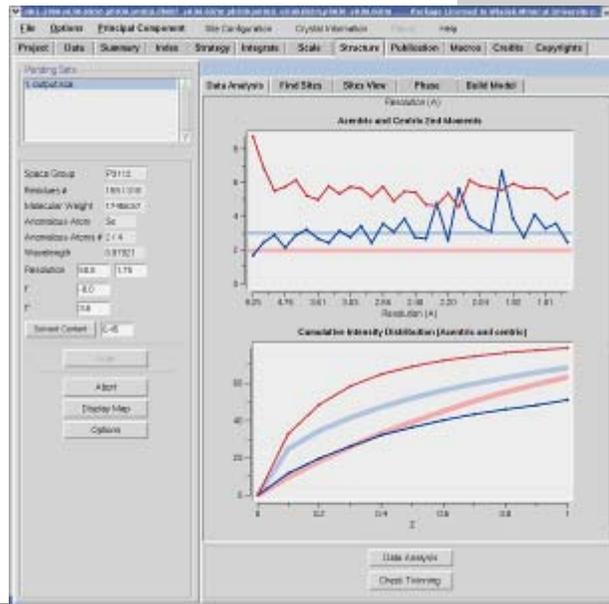
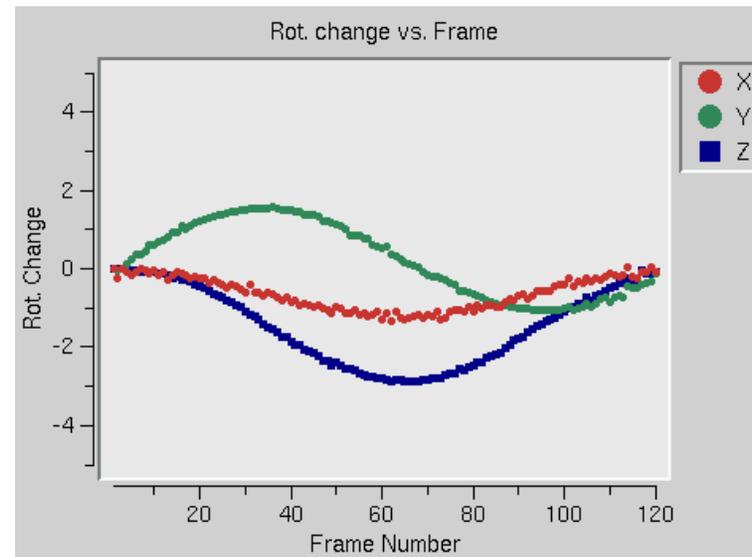
- >80% identity
- 1EJ2 100.000 % global identity in 181 aa overlap
- 1HYB 99.448 % global identity in 181 aa overlap
- >40% and <60% identity
- 1F9A 54.144 % global identity in 170 aa overlap

A 'Protein Sequence' window shows the full amino acid sequence: MMTMRGLLVGRMOPFHRGHLQVIK...GGIS.

A 'Summary <2>' window shows a sequence alignment between the query and PDB entry 1F9. The query sequence is MMTMRGLLVGRMOPFHRGHLQVIK...GGIS. The alignment shows high similarity, with some gaps in the 1F9 sequence (e.g., QSLKDYDIT--). A table below the alignment shows amino acid counts: His (H): 5, Asn (N): 4, Thr (T): 6, Ile (I): 15, Pro (P): 10, Val (V): 15.

Feedback of Possible Hardware, Sample or Experiment Problems

- Detector
- Goniostat
- Cooling
- Crystal decay
- Experiment
- Twinning
-



User Database

- User registration
 - User Biographical Information
 - User Run Information
- Data Collection / Processing Statistics
- Publications
- Calendar
- Forms
- Searches

The screenshot shows a web browser window titled "usercheckin - Microsoft Internet Explorer". The address bar shows "http://get.bio.ars.gov/bsuser/". The page title is "SBC User Registration System (SURS)". The main content area has a green header "SBC Check In" and a form with the following fields:

- ANI Badge Number *
- Last Name
- Birth Date (e.g. January-21-1913) with dropdowns for Month (January), Day, and Year.
- Enter your System Time Start Date: (e.g. January-21-2005) with dropdowns for Month (February), Day (6), and Year (2005).

There is a "Continue" button at the bottom. To the right of the form, there is a text box explaining that the badge number appears on the back of the ANL badge and to the right of the magnetic strip. Below this text is a small image of a badge with a barcode and the text "third through seventh digits of this number".

The intent is to supplement the APS databases, not duplicate.

Upgrades, Upgrades, Upgrades!

- **SBC facility is 10 years old and needs major upgrade - recommended by the APS SAC**
- More than Doubling Capacity of SBC Sector Adding Microbeam Capability
- Upgrading Existing Beamlines to 3rd Generation Technology
- Improving Stability and Reliability of Beam on Sample (R&D)
- Preparing X-ray Optics for 200 mA Beam Current (R&D)
- Fully Integrated Auto-Mounting and Auto-Centering of Samples (R&D)

Improvements will Create new Capabilities at the SBC

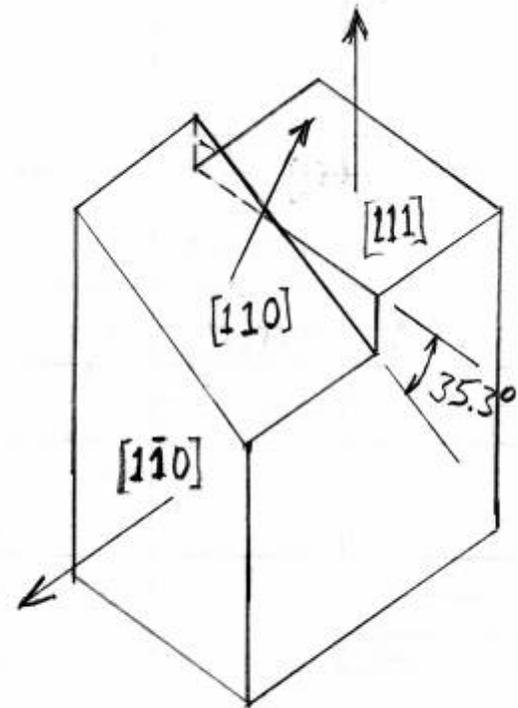
- Improve beam stability
- Automated energy change and tuning
- Robotic handing protein crystals, auto-centering and auto-screening
- Remote data collection capabilities
- Phasing with weak anomalous scatterers (S, P)
- Better design and execution of experiments
- Automated beamline set up
- Automated fluorescence scan

New Development and Upgrade R&D at SBC

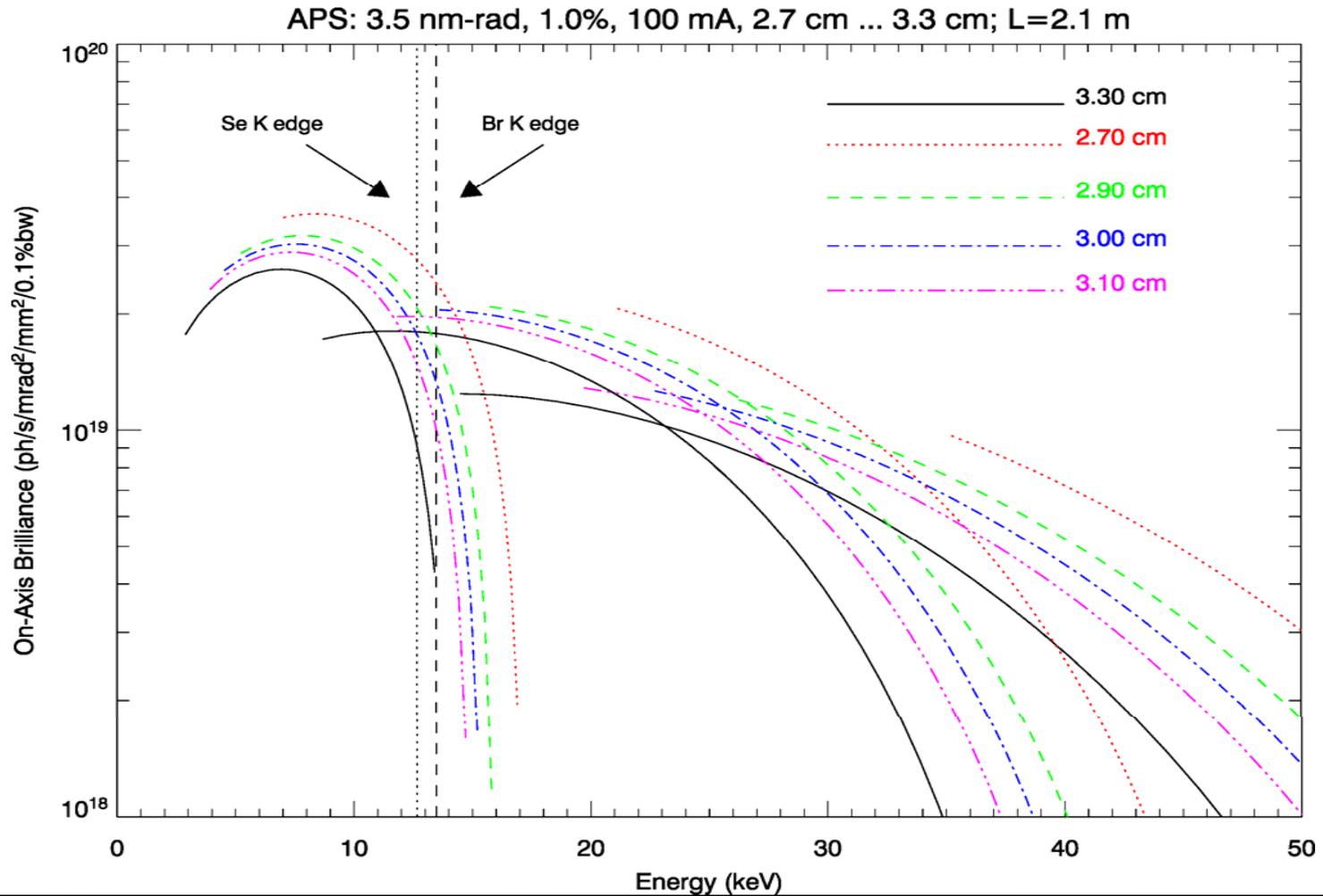
- Hard X-ray Beam Position Monitor
(joint R&D with APS)
- Monochromator Upgrade Design
- Change of Magnetic Period of Undulator
- Need for Microbeam (size, properties)
- Monochromator Upgrades Required for Increase of Beam Current to 200 mA

Dual Cut First Monochromator Crystal (Liquid Nitrogen Cooled for ID)

- Change of energy range by lateral translation of crystal
- No warm-up of cryo-cooled crystal and change of LN2 connections
- Extends energy coverage for limited range of Bragg-angles
- Requires robotic changer of focusing second crystal



Optimize Undulator Source for Biological Applications



Need for Microbeam

- Very small crystals: reduce scatter from non-crystalline material
- Selective exposure of small crystal volumes:
 - very asymmetric shape of crystal (e.g. needle): reduce scatter from non-crystalline material
 - very small, well ordered domains
- Reducing radiation damage in exposed volume:
 - photo-electrons travel several μm ($\sim 6\mu\text{m}$ for 18 keV initial energy)
 - large fraction of damaging energy is not deposited in the illuminated volume for micrometer size beams
 - energy deposit per distance traveled is not uniform, very high at end of travel
- Photo-electrons ejected predominantly in direction of electric field vector

Synchrotrons Productivity: Where are we Now?

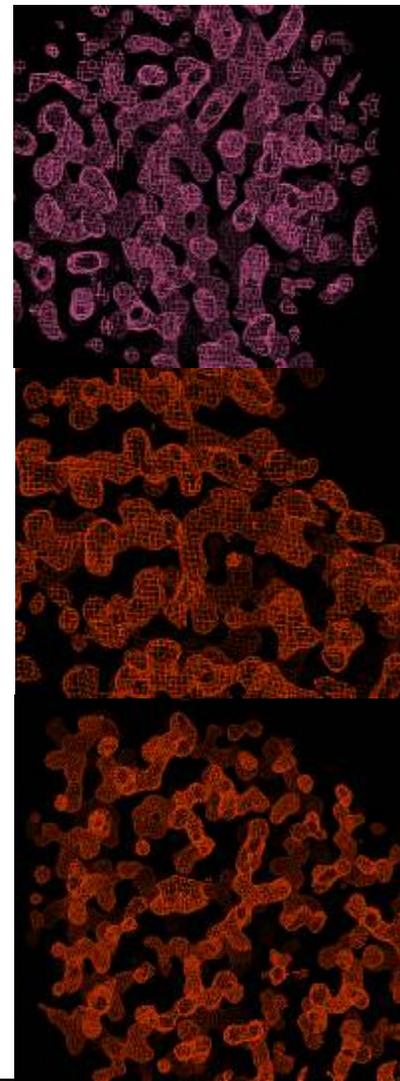
W0308

Teaching Elves to Collect Data: An Analysis of the Last Million Diffraction Images from ALS 8.3.1.
James Holton, Physical Biosciences, Lawrence Berkeley National Laboratory, Berkeley, CA 94720.

Most X-ray data sets collected at synchrotron sources do not produce usable results. An analysis of data collected in 2003 at the ALS beamline 8.3.1 shows that 2346 datasets were collected and 41 structures were deposited in the PDB. Although it is understandable that not every dataset leads to a published structure, it is troubling that ~98% of them do not. This large gap between collected data and useful results is not unique to 8.3.1. The 28 operating American PX beamlines collect ~100,000 datasets/year. This suggests that a great deal of improvement in scientific productivity can be attained if the reasons for failed projects are better understood.

How to Increase Efficiency of Synchrotron Beamlines?

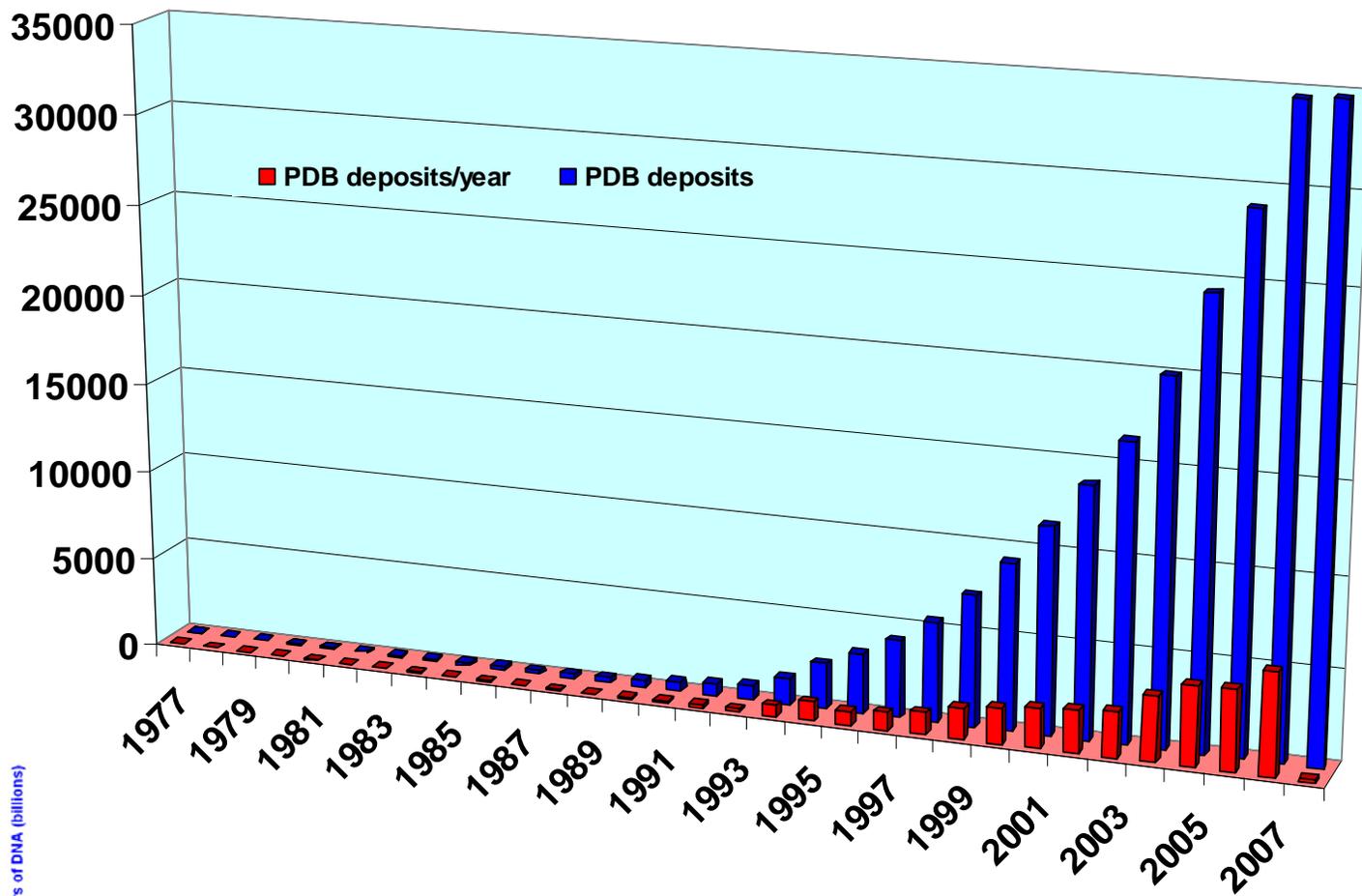
- SBC provides new concept for synchrotron beamlines:
 - **The product of beamline is not diffraction data but interpretable electron density map**
 - Full integration synchrotron hardware, detectors and beamline and crystallographic software
 - Interpretation of initial data, optimization of diffraction experiment and structure solution can be done in real time
 - Crystals can be analyzed at general user facility, but data should be collected at the facility that is best suited for the experiment
 - Multiple experiments on a protein crystal can be performed simultaneously and generate interpretable output in real time
 - Increase in data storage and computing resources
 - Automation and robotics



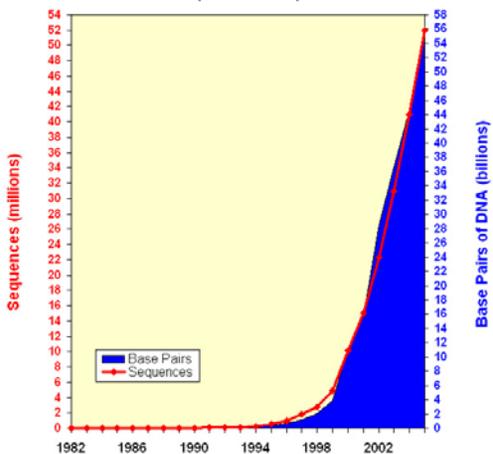
Major Changes in Structural Biology (cont.)

- What is the future – automation and robotics
 - Development of completely automated facilities for data collection and structure determination
 - Hardware, software and procedures – real-time automated sample handling, experimental design, data collection and structure determination and refinement
 - Integration of many resources using IT
 - Remote operations
 - Wide increase in inexpensive HTP technologies in biology – biologists will become major users of synchrotron facilities
- Need for new advanced ID beamlines

Growth of PDB X-ray Deposits



Growth of GenBank
(1982 - 2005)



Advanced Protein Crystallization Facility

- A proposal to the State of Illinois to build the Advanced Protein Crystallization Facility (APCF)
- The APCF is a new \$33M state-of-the-art building
- The building will host advanced crystallization laboratories and user facilities
- The planning funding (\$1.5M) has been awarded this year and we are in the process of assembling the team

SBC and Collaborators

ANL/SBC

R. Alkire
C-s. Chang
M. Cuff
N. Duke
S. Ginell
K. Gofron
Y-c. Kim
K. Lazarski
J. Lazarz
M. Molitsky
J. Osipiuk
T. Petrova
G. Rosenbaum
F. Rotella
K. Tan
R-g. Zhang

Univ. of Virginia

W. Minor
M. Cyborowski
M. Chruszcz
O. Kirillowa

ANL/APS

D. Shu
C. Preissner
D. Nocher
Y. Han
P. Lee
Z. Cai
J. Lee
T. Madden
S. Ross

SER-CAT

J. Gonczy

GM/CA-CAT

B. Fischetti

ANL/HEP

P. DeLurgio

Univ. of Texas SWMC

Z. Otwinowski

S. Park, GS Univ.

Z. Dauter, NIH/NCI

T. Terwilliger, LANL

ANL/BIO

H. Gaines
R. Patel

Funding: DOE and NIH