

# strubix.com

Structural Bioinformatics, Inc. • Behind Every Great Structure

## INSIDE THIS ISSUE

"Access to structural knowledge has led the way to every major commercial revolution."

### Editorial:

#### Structure Enables

Edward T. Maggio, Ph.D.,  
President & CEO, Structural  
Bioinformatics, Inc.  
Page 2

#### Improved Methods for Protein Structure Prediction

Ole Lund, Ph.D., Senior  
Scientist, SBI Advanced  
Technologies A/S  
Page 3

#### Experience and Use of a Reality Center in Drug Discovery

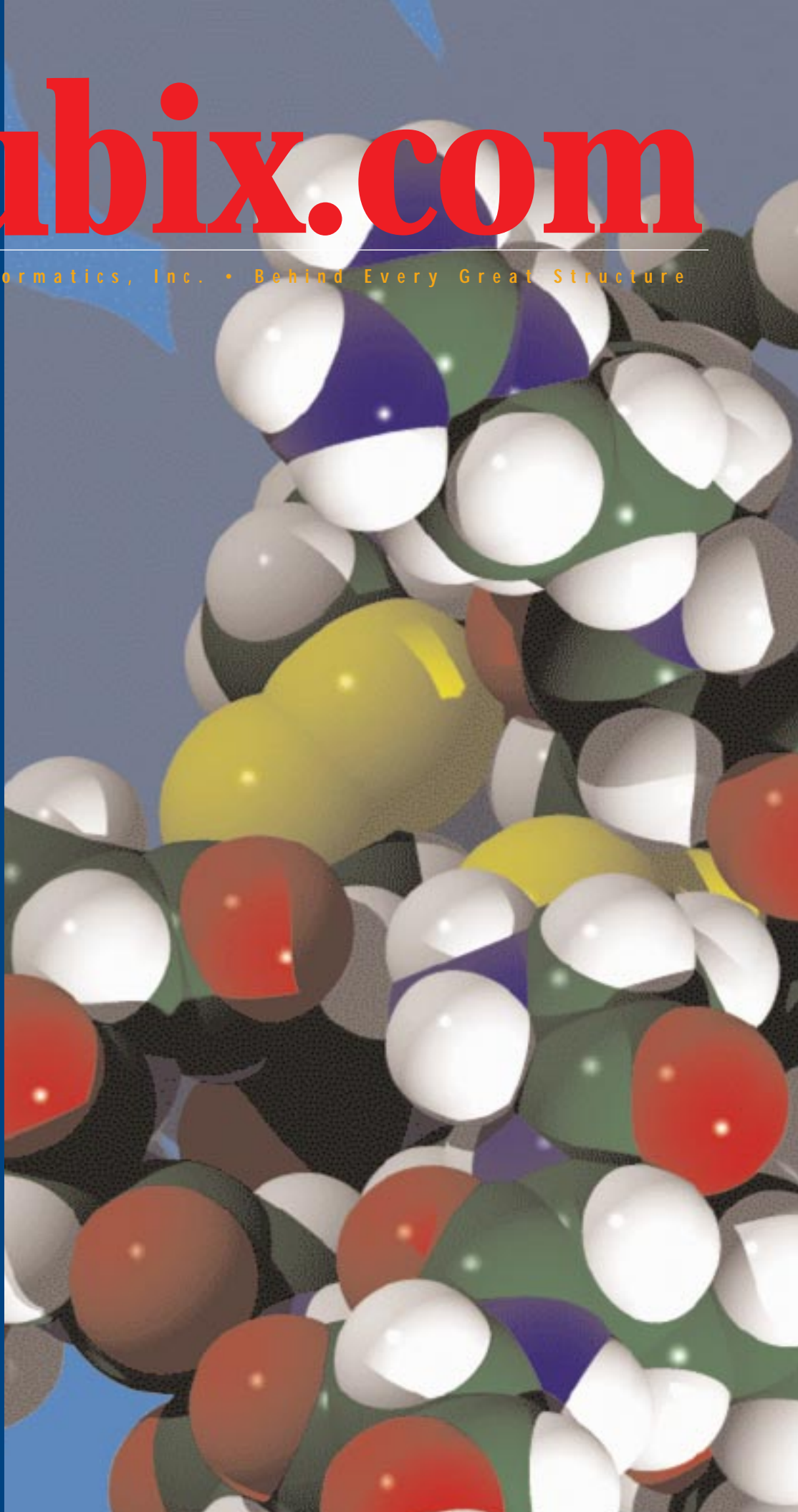
Martin Norin, Ph.D., Director of  
Structural Chemistry, Pharmacia  
Corporation  
Page 4-5

#### SBI Makes a Statement at the Protein Society

Page 5

#### News & Notes

Page 6



# Editorial

## Structure Enables!



**strubix.com**

**A magazine dedicated to the advancements and uses of protein structure for drug discovery.**

Editors:

Alicia Althoff

Ole Wiborg

Frank Mikkelsen

Please direct free subscription orders, inquiries, comments, and questions to:

**In the USA:**

Structural Bioinformatics, Inc.  
10929 Technology Place  
San Diego, CA 92127  
USA

Tel: 858.675.2400

Fax: 858.618.1040

E-mail: [info@strubix.com](mailto:info@strubix.com)

**In Europe:**

SBI Advanced Technologies A/S  
Agern Alle 3  
DK-2970 Hørsholm  
Denmark

Tel: +45.45.16.28.00

Fax: +45.45.16.28.02

E-mail: [info@strubix.com](mailto:info@strubix.com)

Access to structural knowledge has led the way to every major commercial revolution in the chemical and physical sciences over the last two centuries . . . beginning with the understanding of chemical structure in the early 1800's, the elucidation of atomic structure in the early 1900's, the comprehension of polymer structure in the 1920's, and more recently the determination of gene structure in the 1970's. At the beginning of this new century, industrial-scale access and application of 3-D and 4-D protein structural information, across all life science disciplines, promises a scientific and commercial revolution of unequalled impact upon mankind. SBI is catalyzing the revolution in life science R&D by enabling rational experimental design through large-scale protein structure accessibility. We do this through our structural content products, which include SBdBase™ - a steadily growing encyclopedic resource of more than 200 protein families comprising thousands of atomic-resolution protein structures, and SVdBase™ - target-specific structural variant database modules containing tens of thousands of patient-derived protein structures for a growing list of targets such as HIV-Protease, HIV-RT, and many others. We also do this through our atomic-resolution custom modeling service for public or proprietary sequences and through our X-ray crystallographic structure determination services.

Because access to high quality protein structural information fundamentally enables rational experimental design in essentially all life science disciplines -- from molecular biology and pharmacology, medicinal chemistry, and structure based drug design, to the many sub-disciplines of computational biology as a whole -- the major "Post Genomic Era" challenge will be to translate all of this critically valuable sequence data into structural knowledge.

The fact that structure will be the key to unlocking the coming post-genomic revolution in human medicine is confirmed by the scientific community's commitment to a

growing number of large-scale "structural proteomics" projects in the United States, Europe, and Japan. Most of these efforts are based on attempts to scale-up physical methods such as X-ray crystallography. However, the principal hurdles of cloning, expression, purification, and crystallization remain idiosyncratic processes, not easily subject to acceleration. In addition, these processes account for the fact that only one in 20 proteins ultimately yields useful crystals, while some proteins, such as membrane-associated proteins, may not crystallize at all. Nevertheless, these activities will make important contributions in expanding the scope of augmented homology modeling and other advanced modeling methods. It is in these latter fields that SBI has its core competence.

Through intensive technology development efforts at our Danish subsidiary SBI Advanced Technologies (SBI-AT), our Cambridge, Massachusetts subsidiary (SBI-Moldyn), and at our San Diego headquarters, we strive to be the world's leading provider of large-scale, highest quality protein structural databases and advanced structural data mining technologies. As one example, we invite you to look at the article by Dr. Ole Lund in this current issue of *strubix.com*. Dr. Lund describes how SBI has expanded the borders of reliable structure prediction using sophisticated neural networks by incorporating advanced "Wring Mode" concepts developed at SBI-AT to produce the most accurate secondary structure predictions currently available from any source. In coming issues, we'll describe additional advances in the acceleration of X-ray and NMR structure solution, long time-frame dynamics, novel multi-variate analysis technology for *ab initio* identification of function from 3-D protein structures, and more.

Edward T. Maggio, Ph.D.  
President & CEO





# Improved Methods for Protein Structure Prediction

By Ole Lund, Ph.D.

The team of scientists working at SBI Advanced Technologies A/S (SBI-AT) is developing novel technologies for protein structure prediction. The goal of this work is to be able to make accurate tertiary structure models for as many protein sequences as possible.

The work covers diverse areas such as prediction of protein secondary structure using Neural Networks, construction of improved sequence profiles, Hidden Markov Models using sequence and structure profiles, construction of non-redundant data sets, and construction of novel force fields. An algorithm for predicting secondary structure of proteins at 80% will be published in a forthcoming issue of *Proteins* [1].

The ability to predict protein secondary structure accurately is a step towards the ultimate goal of being able to predict the tertiary structure of proteins from the amino acid sequence.

Accurate secondary structure predictions significantly enhance the capability to make accurate protein models. The secondary structure predictions can be used to recognize folds and find templates for remote homology modeling by identifying other proteins with the same composition and sequential order of secondary structure units. It can also be used to increase alignment accuracy, as well as aid in finding fragments for *ab initio* structure prediction.

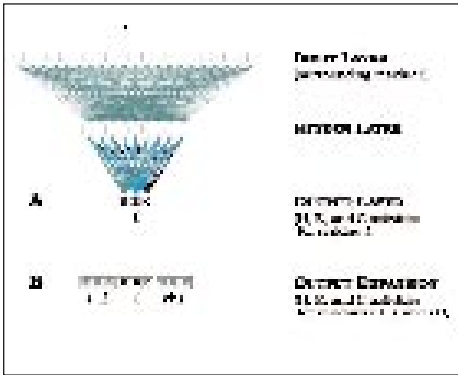
The secondary structure predictions will enable SBI to make more accurate pro-

tein models and create models for proteins that were previously too different from any known protein structure. In turn, this will facilitate the search for potential drugs that will bind to these proteins.

Use of up to 800 predictions of differently trained Neural Networks, and the ability to combine the networks in an efficient manner, lead to a more accurate prediction than that of any of the individual networks. A novel technique: output expansion (Figure 1) that predicts the secondary structure for more than one residue at a time is also a key element of the new method. This improves the prediction accuracy by teaching the neural network about the structural context of its secondary structure predictions. The method not only calculates the most likely secondary structure for a given residue, but also calculates the probability that a residue is in any of the three secondary structure conformations. This type of output is much more useful as input to probabilistic methods such as Hidden Markov Models.

Using these new technologies the secondary structure of proteins can be predicted with an unprecedented 80% accuracy rate, thus improving the state-of-the-art in this very competitive field.

Dr. Lund is a Senior Scientist at Structural Bioinformatics Advanced Technologies A/S (SBI-AT). SBI-AT is located in Hørsholm just north of Copenhagen, Denmark. [1] Prediction of protein secondary structure at 80% accuracy. *Proteins*, in press, 2000.

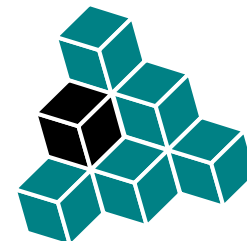


**Figure 1:** Schematic drawing of a neural network for predicting protein secondary structure. The neural network reads the amino acid sequence into the input layer. By use of a hidden layer in the neural network, the input is transformed in a nonlinear way into a secondary structure prediction. The probability for alpha-helix (H), beta-sheet (E), and coil (C) for one residue is calculated in the output layer of the neural network (A). When using output expansion, the secondary structure is calculated for three consecutive residues in a single step (B).

Tomas Nordahl Petersen, Claus Lundegaard, Morten Nielsen, Henrik Bohr, Jakob Bohr, Søren Brunak, Garry P. Gippert and Ole Lund.



# Experience and Use of a Reality Center in Drug Discovery



By Martin Norin, Ph.D.

*In Pharmacia Corporation we have at our research site in Stockholm, since one and a half year, in-house access to a Reality Center. The main purpose for the investment was to be able to share 3-D structural data among members of project teams. Our experience is that the impact of visualizing 3-D data using the Reality Center is important. So far the Reality Center has been mostly used in medicinal chemistry projects. However, we see a growing interest from target validation teams. Based on our experiences so far, there are a couple of items to address when planning a Reality Center. The use of the Reality Center has increased the overall understanding and use of structural data in our research.*

Protein 3-D structural data has become more and more important in our genomics-driven research projects. The uses of structures, both high resolution experimental data and low resolution protein models, in decision making now also go beyond what we call 'Structure Based Iterative Design'. In the early exploratory research, many targets are picked up by genomic-based approaches. These targets need to be validated and prioritized before substantial investments are made to form full-blown medicinal chemistry projects. Structural data can here provide very useful insights into the 'drugability' of a target (the ability for small molecules to bind to relevant target sites), opportunities for selectivity, design of the screening funnel, and the prediction of pharmacogenomic effects (structural mapping of SNPs).

The motivation for the investment in a Reality Center was a need to enable project teams to collectively and interactively share and discuss 3-D structural data. The 'Reality Center' at our research site is basically a high performance SGI Onyx computer graphics



system coupled to a large screen projection system placed in a regular conference room. It requires the use of passive polarized 3-D glasses, and up to 17 people can easily share interactive 3-D views. Our experience so far is that the use of this facility has been important for the impact of structure-based design on medicinal chemistry teams. Structural and computational chemists are very well aware of how

important and sometimes also how difficult it is to get synthetic medicinal chemists involved in the 'rational' design of new molecules. Here, we have experienced that the use of the 'Reality Center' in chemical library design sessions has really unlocked the potential and creativity of the involved structural and medicinal chemists. In addition, as mentioned above, structural data has also become important in other areas of drug discovery. Here, we have seen that the 'Reality Center' boosts the impact of structural data on project and research management.

Our facility has also provided us with a few bonus values that we did not anticipate before the investment. The Reality Center is an illustrative and impressive tool for presentations to external visitors. In addition, it has been very useful for our long-term recruitment purposes. Molecular graphics sessions held at the Reality Center are very appreciated during site visits by students of universities and other schools.

We have experienced some practical





A montage illustrating the 'Reality Center'. The two scientists, Anna-Lena Gustavsson and Mats Kihlén, are wearing the rather convenient 3-D glasses. The large screen is in the background. The blue molecular surfaces are mimicking the impression provided by the system when sitting in the room having three-dimensional molecular models 'floating' around in the room.

challenges that may be useful to consider when planning for a 'Reality Center'. The location should be easily accessible and centrally located to facilitate spontaneous use. Try to avoid competition from other types of meetings, and at least do not co-locate it with a video conferencing facility! To save space we have chosen a front projection system. It works well although the room needs to be darkened to provide enough intensity. A back projection system is said to have somewhat higher intensity, but such a system requires a large space (several meters) that is only used to project the picture on the screen. Make sure that the hardware 3-D graphics of your favorite graphics software is compatible with upgrades of the computer system. It is also very effective to create an easy and user friendly IT infrastructure. In addition to computer graphics software, this involves clear and consistent project areas for the data, pre-prepared 'project views' with annotated

structural data, and the possibility for bench scientists to continue by their own to explore the data with the same interactive views on their own office PC's. There are only a few molecular modeling software packages that enable these functions. We are making extensive use of an in-house software which we call Mosaic2. This software enables all functions mentioned above, and it is also interfaced with our corporate cheminformatics system. The program also enables 'tele-modeling' sessions. This function is primarily used to exchange data during 'site-to-site' modeling sessions between individual scientists. However, on a few occasions we have had shared interactive trans-Atlantic sessions between our 'Reality Center' in Sweden and a similar facility at our research site in Kalamazoo, Michigan, USA. During such a session, I really felt that we are now into the era of being part of a globally connected information technology environment!



**Dr. Martin Norin**  
**Director of Structural Chemistry**  
**Pharmacia Corporation**  
**Discovery Research, Stockholm**

## The Fourteenth Symposium of the Protein Society San Diego, CA August 5-9, 2000

Structural Bioinformatics made a commanding appearance at this year's symposium of the Protein Society by demonstrating its knowledge and cutting-edge technology in the field of protein modeling. In addition to having an exhibition booth at the meeting, the following scientific posters were presented:

1. Molecular Modeling Study of Drug-Resistant Mutations in HIV-1 Protease - Inhibitor Complexes. **M.D. Shenderovich, V.M. Tseitin, C.L. Fisher, J. Kottalam and Kal Ramnarayan**
2. Molecular Modeling and Analysis of TNF-alpha/TNFR and CD40/CD40L Complex. **G. Raghunathan and Kal Ramnarayan**
3. Comparison of Protein Homology Models from Structural Bioinformatics' Database (SBdBase™) to Model Structures from Public Databases Created by Fully Automated Comparative Modeling Methods. **Michael J. Dudek, M. Prabhakaran and Kal Ramnarayan**
4. Fast Scanning of Protein Database for Identification of Receptor Binding Sites and Evaluation of Binding Affinity of Ligands. **M. Prabhakaran and Kal Ramnarayan**
5. Sequence, Core and Surface Similarities within Proteins in a Large Database: SBdBase™, a Case Study. **Saied Moezzi, Behnam Vessal, M. Dudek, Suresh Tudor, M. Prabhakaran and Kal Ramnarayan**
6. Augmented Homology Modeling of the Human Her2 (erbB2) Protein. **Jianhua Zheng, Christina C. Niemeyer, Michelle N. Endo, Seymour Mong and Kal Ramnarayan**
7. Homology Model of the beta-Secretase BACE protein, a Prime Target for Alzheimer's Disease Drug Development. **Shankari E. Mylvaganam and Kal Ramnarayan**

For further information or to obtain reprints of the abstracts, please visit our website: [www.strubix.com](http://www.strubix.com). You may also send e-mail requests to: [info@strubix.com](mailto:info@strubix.com).

All intellectual property is owned by Structural Bioinformatics, Inc., 2000®.





## SBI COMPLETES \$32.6 MILLION FINANCING

On March 31, 2000, SBI successfully closed a \$32.6 million Series D preferred stock financing. Due to substantial over-subscription, the total Offering was expanded from an initial goal of \$15 million to \$32.6 million. "The response to SBI's business focus has been outstanding," said Dr. Edward T. Maggio, Chairman, President and Chief Executive Officer of SBI. "As sequencing of the human genome is nearing completion, we are seeing a fundamental evolutionary shift in attention toward genome-derived protein structure and structure-based drug design. SBI is positioned to become the main source of highest quality, large-scale, proprietary 3-D protein structure and structural polymorphism information for the decade after the genome." "Also, it should be noted that the added proprietary information content generated by SBI as it converts gene sequences to 3-D protein structures constitutes a very real and viable response for genomics companies concerned not only about preserving, but actually significantly expanding, the value of their proprietary intellectual property assets."

The additional funding received in March has allowed for expansion of the Company's computational science, drug discovery science and database infrastructure development. Construction has been a familiar sight in San Diego during the spring and summer months as SBI expanded its San Diego facilities to include a building adjacent to the headquarters. The additional office is located at 10949 Technology Place and houses the Accounting, Corporate Development, Sales & Marketing and Human Resources Departments. With the departure of these groups from the main building, construction on the expanded Chemistry facility has begun.

## QUALITY OF MODELED STRUCTURES IN THE SBdBASE™ CONFIRMED BY X-RAY CRYSTALLOGRAPHY

Structures for 30 of the proteins modeled in the SBdBase™ were recently solved by X-ray crystallography. A comparison of the 3-D coordinates reveals a remarkably accurate agreement between the predicted models and their crystal structures. "The new data verify the very high quality of the protein models in the SBdBase™," said Ole Wiborg, Managing Director of SBI-Advanced Technologies A/S. "The results strongly indicate that protein structures built using SBI methodologies are as valuable as structures solved by NMR or X-ray technologies."

*A detailed analysis of the results from the comparison study will be published in the next issue of strubix.com.*

## PROTEIN STRUCTURE AND VISUALIZATION SEMINAR

On May 23, 2000, SBI Advanced Technologies, Silicon Graphics and The Danish Computing Center for Research and Education (UNI·C), held a seminar on new methods in the areas of structure/function prediction and 3-D visualization of proteins. The seminar took place at the newly opened Virtual Reality Center located on the premises of the Technical University of Denmark in Copenhagen. In the center's conference and VR demonstration facilities, applications of protein structure information and the use of graphical display tools in drug discovery were presented and discussed by leading scientists in the field. One of the many informative and interesting presentations, entitled "Experience and Use of the Reality Center in Drug Discovery," was given by Dr. Martin Norin, Director of Structural Chemistry for



Pharmacia Corporation (Stockholm, Sweden).  
*The article written by Dr. Norin in this issue of strubix.com is based on his presentation during the seminar.*

## SBdBASE™ VERSION 2.2.2 NOW AVAILABLE

SBI's comprehensive database of protein structures is now available in a new and improved 2.2.2.version. The SBdBase™ is a relational database of 3-D and 4-D protein structures, which promises to accelerate lead discovery and optimization and to enable rational experimental design across broad disciplines within the pharmaceutical and life science industries. The new version contains 2,500 structures representing more than 200 protein families. The new features include:

- Individual and simultaneous viewing of multiple proteins
- Interactive viewing of sequence and structure inquiries
- Surface mapping of electrostatics and structure dynamics
- Comprehensive quality assurance plots
- Import/export structures for comparative purposes
- Advanced structural data mining and analysis tools



# Structure Enables!

Access to Structural Knowledge Has Led the Way to Every Major Commercial Revolution in the Chemical and Physical Sciences Over the Last Two Centuries



... beginning with the understanding of chemical structure in the early nineteenth century, the elucidation of atomic structure in the early 20th century,



and more recently the determination of DNA and gene structure. At the beginning of this new



century, industrial-scale access and application of 3-D and 4-D protein structural information across all life science disciplines, promises a scientific and commercial revolution of unequaled impact upon mankind.

**SBI is revolutionizing life science R&D by enabling rational experimental design through large-scale protein structure accessibility.**

## Our structural content products include:

**SBdBase™** - a steadily growing encyclopedic resource of more than 200 protein families comprising thousands of atomic-resolution protein structures

**SVdBase™** - target-specific modules containing tens of thousands of patient-derived protein structures for a growing list of targets such as HIV-Protease, HIV-RT, HIV-Integrase and many others

**CombiLib™** - virtual combinatorial library modules built upon pharmaceutically-attractive chemical scaffolds

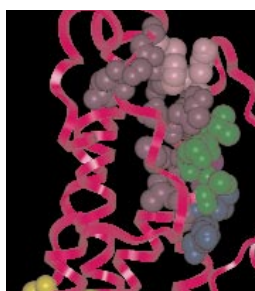
**ProteoMine™** - an integrated software system for sequence and structure analysis and automated data extraction, filtering, and report generation from hundreds of internet accessible data sources

## Our computational drug discovery collaborations provide:

- A 1000-fold reduction in the number of compounds acquired and tested
- Lead molecule generation from gene sequence in as little as 60 days
- Multiple chemical scaffold output for broad patent coverage and back-up series

## Our structural services include:

- Atomic-resolution custom 3-D modeling of public or proprietary protein sequences
- X-ray Crystallographic Structure Determination



**For molecular biology, medicinal & computational chemistry, protein modeling, pharmacogenomics or high throughput screening, Structure Enables!**



**Call SBI today to see how structure can enable your success.**

**Structural Bioinformatics, Inc.** 10929 Technology Place, San Diego, CA 92127

E-mail: [info@strubix.com](mailto:info@strubix.com) • Tel: 858-675-2400 • Fax: 858-618-1040 • Website: [www.strubix.com](http://www.strubix.com)

# Upcoming C o n f e r e n c e s

## Europe

### Biotechnology 2000

September 3-8,  
Berlin, Germany

http:

//dechema.de/englisch/veranst/ibs11/pages/ibs11\_1.htm

### Medicon Valley Bioconference 2000

September 20-21,  
Malmö, Sweden

http:

//www.mva.org

### Biomics Gene to Protein, Protein to Structure, Structure to Drug

November 13-16,  
Stuttgart, Germany

http:

//www.ibr-biomics.com

### Bio-Europa 2000

November 13-15,  
Berlin, Germany

http:

//www.ebdgroup.com/CONNECT

## Biomedical Partnership Forum

November 28,  
Göteborg, Sweden

http:

//www.sbic.se/

## USA

### The 14th Symposium of the Protein Society

August 5-9,  
San Diego, CA

http:

//www.faseb.org/meetings/protein00

### IBC's 5th Annual World Congress Drug Discovery Technology 2000

August 14-16,  
Boston, MA

http:

//www.drugdisc.com

### IBC's Annual Symposium on Proteomics, Functional Genomics & Bioinformatics

October 16-19,  
Phoenix, AZ

http:

//www.ibcusa.com

### CHI's Second Annual Protein Structure

October 26-27,  
McLean, VA

http:

//www.healthtech.com/conference/00pst/index.htm

### UCSD CONNECT

### Corporate Partnership Forum/CALBIOsummit

October 30-31,  
San Diego, CA

http:

//www.calbiosummit.org

### CHI's Second Annual Research Informatics

November 29-30,  
Philadelphia, PA

http:

//www.healthtech.com/conference/00dmn/index.htm

REPRESENTATIVES FROM  
SBI AND SBI-AT WILL BE PRESENT  
AT ALL OF THESE EVENTS AS A SPEAKER,  
EXHIBITOR, OR BOTH. CONTACT US FOR  
MORE INFORMATION.



STRUCTURAL BIOINFORMATICS, INC.  
10929 TECHNOLOGY PLACE  
SAN DIEGO, CALIFORNIA 92127

TEL: 858.675.2400  
FAX: 858.618.1040  
WWW.STRUBIX.COM

SBI ADVANCED TECHNOLOGIES A/S  
AGERN ALLÉ 3  
DK-2970 HØRSHOLM  
DENMARK

TEL: +45.45.16.28.00  
FAX: +45.45.16.28.02



## BULK RATE

U.S. Postage Paid  
Permit #760  
San Diego, CA