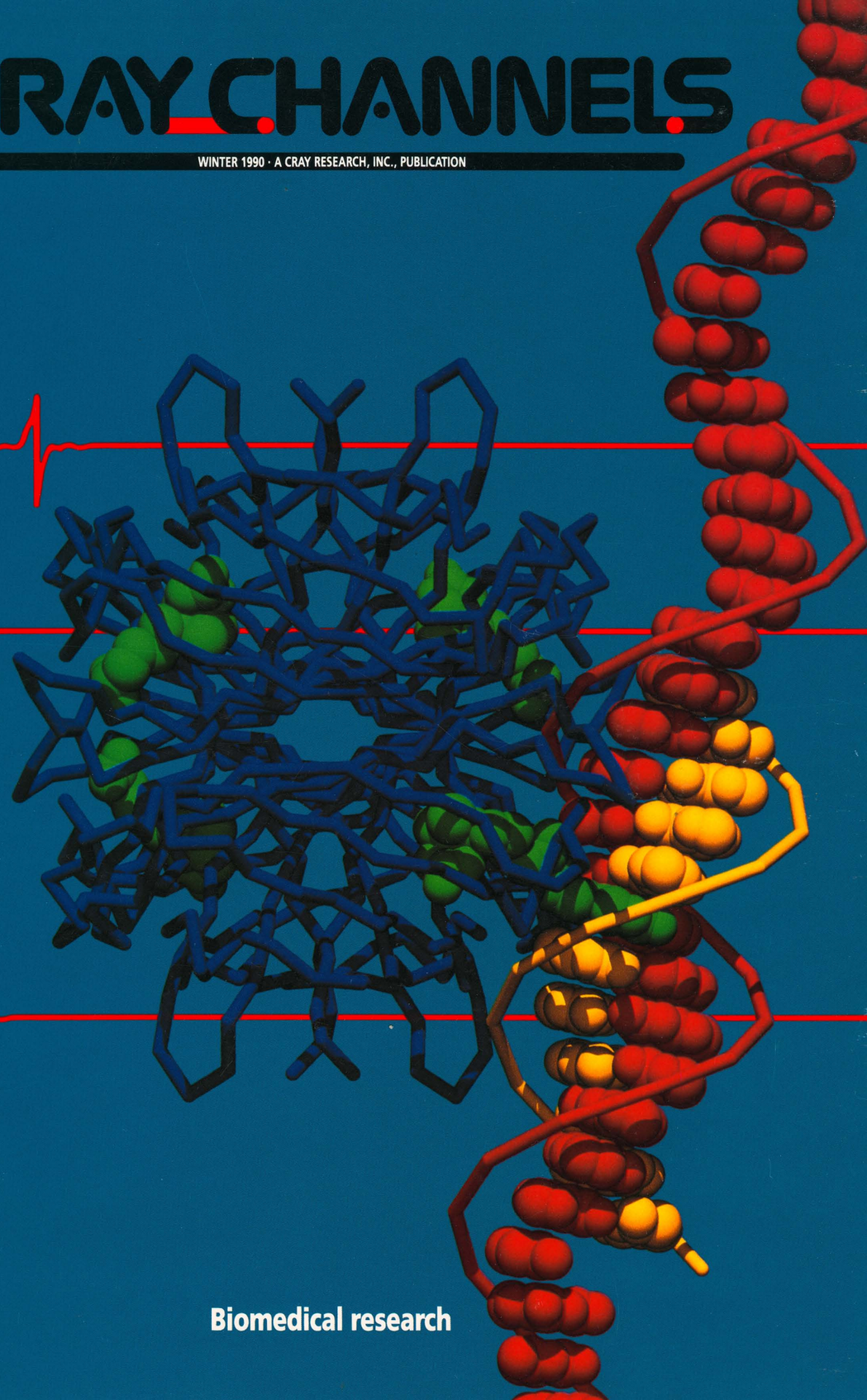
CRAY RESEARCH, INC.
Marketing Communications
1333 Northland Drive
Mendota Heights, MN 55120
612/681-3438

CRAY CHANNELS

WINTER 1990 · A CRAY RESEARCH, INC., PUBLICATION



Biomedical research

CRAY CHANNELS

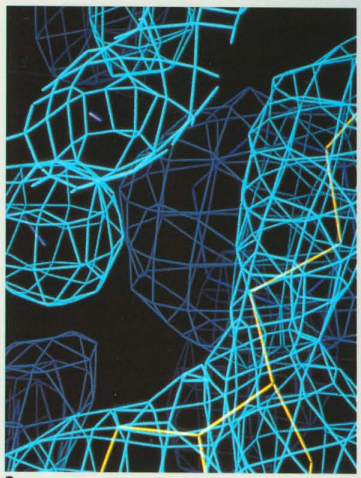
In this issue

By using supercomputers to explore the chemistry of life, biochemical researchers are opening new doors to improved health care. And researchers working in other areas of biomedicine are applying supercomputers to the design of prosthetic devices and to etiological research. In this issue of CRAY CHANNELS, we profile biomedical applications of supercomputers in commercial, university, and government research laboratories.

Researchers at the Du Pont and G. D. Searle companies are using computer modeling to study the biological activities of macromolecules. At the Research Institute of Scripps Clinic, researchers are using supercomputer technology to determine protein structures from nuclear magnetic resonance (NMR) spectroscopy. Researchers at The Ohio State University are using magnetic resonance imaging (MRI) technology to help chart the development of atherosclerosis. NASA researchers have turned their expertise in computational fluid dynamics toward improving the design of artificial hearts. Our regular departments describe enhancements to the Cray Ada Environment, along with a newly optimized convolution routine and profiles of Cray Research's Gigaflop Performance Competition winners.

As the complexity of biological systems yields to the computational power of supercomputers, medical and biochemical researchers are able to enjoy the cost- and time-saving benefits of large-scale computation. Everyone stands to benefit as this research leads to improved health care treatments and preventatives.

Features



2

Departments

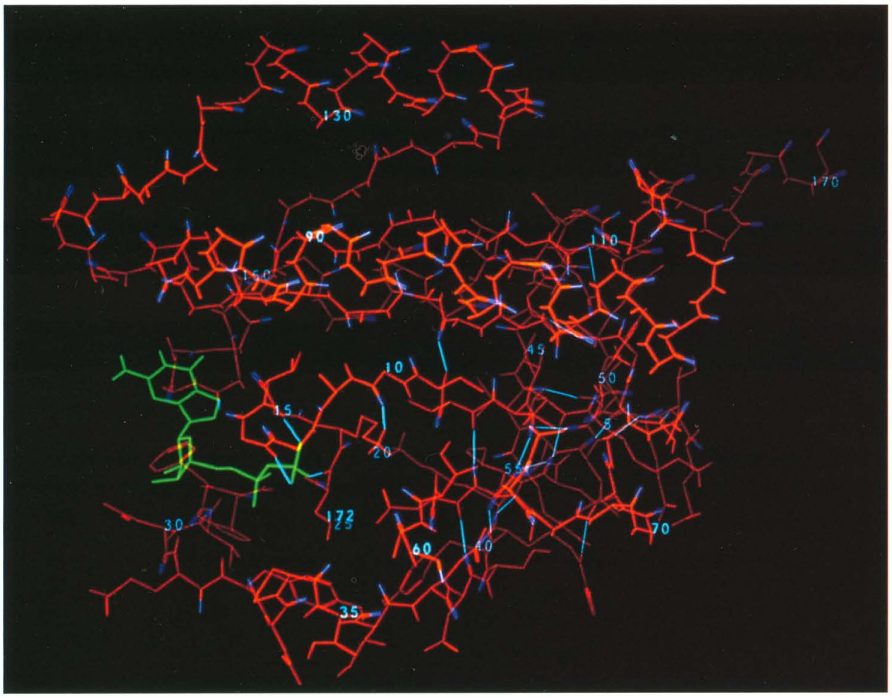
CRAY CHANNELS is a quarterly publication of Cray Research, Inc., intended for users of Cray Research computer systems and others interested in the company and its products. Please mail feature story ideas, news items, and Gallery submissions to CRAY CHANNELS at Cray Research, Inc., 1333 Northland Drive, Mendota Heights, Minnesota 55120.

Volume 11, Number 4

Editorial staff
Ken Jopp, editor
Elizabeth Knoll, associate editor

Design and production
Barbara Cahlander
Eric Hanson
James Morgan
Cynthia Rykken

On the cover: Streptavidin-biotin interactions form the basis for many diagnostic methods that require the formation of a specific and irreversible linkage. The figure shows a backbone representation of the streptavidin tetramer with one bound biotin covalently linked to a single-strand DNA probe molecule, bound in turn to a complementary strand of a larger DNA fragment. Streptavidin graphic output using ANIMOL by Richard Hilmer, Du Pont Central Research and Development Department.



Computational challenges in structure-based drug design

John J. Wendoloski and F. Ray Salemme
E. I. du Pont de Nemours and Company
Wilmington, Delaware

Figure 1. Model of the ras p21 protein showing guanosine diphosphate bound in the protein active site.

An emerging understanding of the structural basis of biological function has made possible a new technology of rational molecular design. An area receiving intense attention is the rational design of new drugs and herbicides that regulate the activity of biological macromolecules. Typically, these molecules bind specifically to catalytic enzymes, cellular receptors, or nucleic acids that underlie physiological processes in living organisms. Effective rational design programs use a combination of modeling and intensive computational tools, but still rely critically on direct structural measurements for verification and direction. Both structural determination and rational drug design require using large amounts of supercomputer time. Much of the work described in this article was performed on the CRAY X-MP/28 computer system at the Du Pont Experimental Station.

Strategies for drug discovery

An established approach to pharmaceutical discovery uses screening methods to test a library of thousands of unrelated molecules in a biological assay system. Traditionally, these assays range from measures of whole animal behavior to assays based on microorganism survival. Lead compounds showing activity in assay screens are "analoged" by systematic chemical modification to produce new rounds of compounds. Statistical and structural comparisons of data from a family of compounds allow inferences to be made about which chemical and structural properties are important in generating the pharmacological response. Deductions based on such quantitative structure-activity relationships (QSAR) can drive synthesis of new leads or new rounds of analogs with improved properties. In a fundamental sense, these methods are aimed at generating a ghost image of the receptor site at which the drug acts, without knowing anything else about it.

Screening remains effective in drug discovery, particularly when a detailed understanding of the molecular basis of a disease allows identification and assay of a specific target molecule. At the same time, the more detailed characterization that usually follows enzyme or receptor isolation provides important information to direct the selection of candidate drugs. In cases where an enzyme is targeted for inhibition, a basic knowledge of the chemistry catalyzed can provide information sufficient to design useful compounds. These can be "transition state analogs," which bind to the enzyme active site more tightly than natural substrates and competitively inhibit activity, or "suicide substrates" that irreversibly react with the enzyme's catalytic machinery. A potential difficulty with a mechanism-based approach is that many enzymes share catalytic features, so that inhibitors generated in this way usually have to be analoged to achieve the required level of specificity. Nevertheless, this is an important approach that continues to provide new drugs and herbicides.

More recently, methods have been developed that attempt to design drug molecules from first principles, based on detailed structural knowledge of binding sites on an enzyme or cellular receptor. Although proteins are heteropolymers of amino acids whose chemical structures are inferred easily from the sequence of the encoding DNA molecule, they fold to form complex three-dimensional structures that presently are impossible to predict. However, by using x-ray crystallographic methods, the three-dimensional structure of a protein can be experimentally determined, providing a view of the enzyme or receptor binding site in atomic detail. As an example, Figure 1 shows the structure of the ras p21 protein, an enzyme whose genetically mutated form is a contributing factor in the growth of some cancers.[1] In principle, detailed structural information should be sufficient to design a drug to bind to the target molecule, but in fact this task is far from simple.

Binding sites on proteins usually form a pocket lined with functional groups that make specific

interactions with the normal substrate or effector. The objective of structure-based drug design is to generate, or alternatively, to retrieve from a structural data base, a small molecule ligand that will fit the binding site and make specific interactions with the protein. At present, no bona fide examples of the success of this approach have been reported that were not heavily based on pre-existing structural precedents. Indeed, the key to effective programs in structure-based drug design involves an integration of the methods of screening, QSAR analysis, data base searches, and computational approaches with iterative x-ray crystallographic studies that allow direct inspection and verification of the ways in which inhibitors bind to the target protein molecule.

Computational tasks in structure-based drug design

Figure 2 illustrates the information flow for a structure-based drug design program. This scheme includes inputs from empirical screens and molecular modeling studies. However, its distinctive feature is the use of x-ray crystallography to directly visualize protein-drug interactions as the definitive means of devising a rational strategy for modification. Many of the stages of the process outlined are computationally intensive. These include searches of structural data bases, molecular modeling studies to design new compounds, and particularly, the iterative computations required for crystal structure determinations of target proteins and inhibitor complexes.

Protein crystallography

A detailed description of the theory and practice of protein crystallography lies outside the scope of this article. However, the basic objective is to induce the molecules of interest to form a regular, three-dimensional lattice. The regularly arrayed molecules in a crystal lattice (Figure 3) scatter x-rays to produce a diffraction pattern that samples the spatial Fourier transform of the individual protein molecules. Structure computation basically involves mathematical inversion of the molecular transform to produce an electron density map of the protein.

Typically, x-ray data are collected as sets of multiple, two-dimensional digital arrays. Raw data for a protein structure determination easily can run

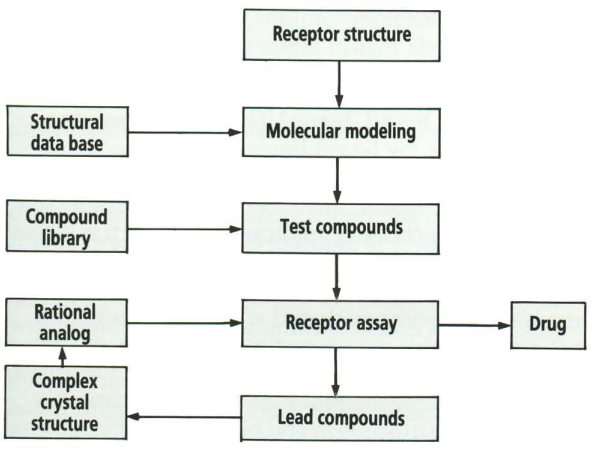


Figure 2. Information flow in a structure-based drug design program.

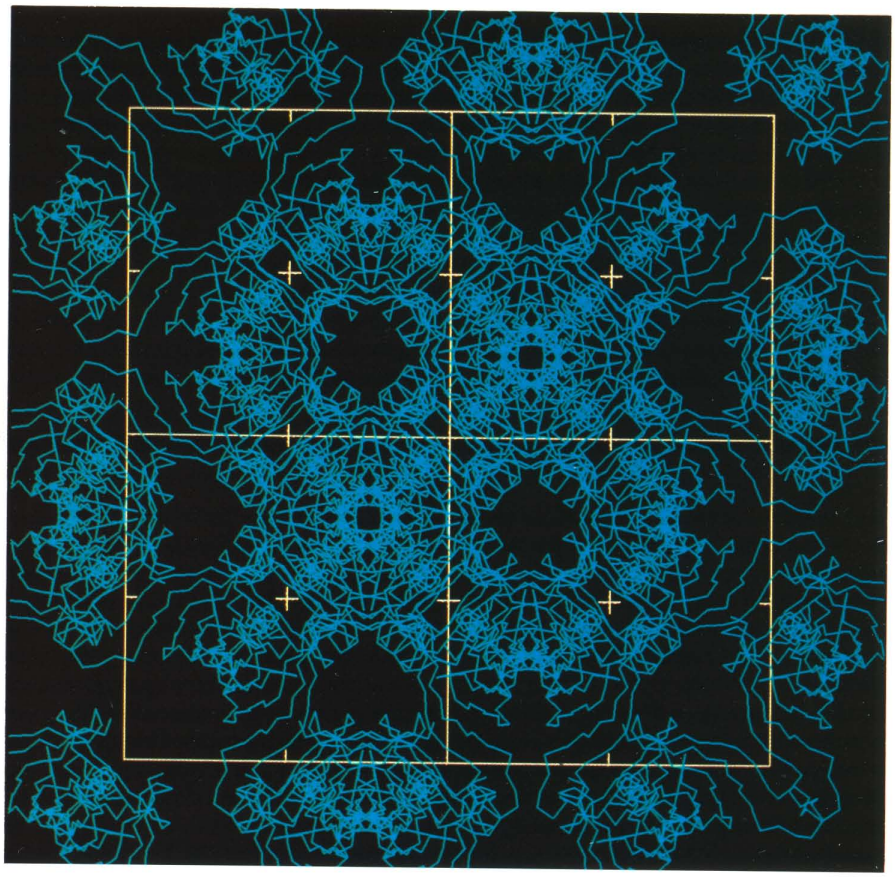


Figure 3. Molecular packing in the streptavidin crystal lattice.

to 2 or more Gbytes for a parent structure and about 500 Mbytes for each drug-complex data set. Because only the power spectrum of the spatial molecular transform can be recorded directly, the missing phase information required to reconstruct the scatterer must be recovered independently. One approach, the method of isomorphous replacement, involves collecting data from crystals of the native protein that have been doped with metals. The simple structure of the doping metal, embedded within the more complex crystal structure of the protein, can be solved by direct methods used for small molecule crystal structures. Once complete, the "heavy atom" structure provides an internal phase reference that can be used to phase the undoped protein crystal structure. Although the computations involved are substantial, they seldom constitute a rate-limiting step with modern computing equipment.

An alternative approach, useful when the structure being determined is similar to one known already but which crystallizes in a different space group, is the method of molecular replacement. Here the structure of the known molecule is used to compute a diffraction pattern that can be compared with that measured from the new protein crystal. If a correlation can be found, then phases provided from the input model can be combined with measured data to reconstruct the new structure. Finding a correlation can be an exceedingly large computational task, as the success of the method is exquisitely sensitive to case-specific details that must be investigated systematically.

Whatever the method of phase determination, inversion of the complex transform produces an electron density map of the protein crystal. The map can be interpreted using computer graphics tools that

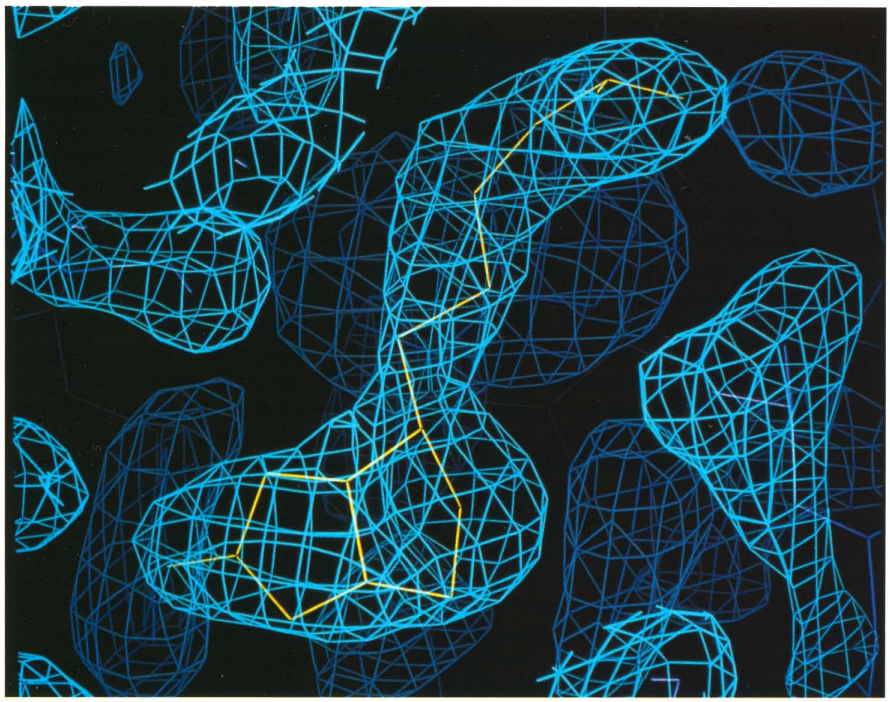


Figure 4. An electron-density map showing the biotin bound to its binding site in streptavidin.

make it possible to fit a stick-bond model interactively with complete conformational flexibility into the electron density (Figure 4). This initial model, which usually has only modest resolution, can be extended and improved using methods of crystallographic refinement. This process involves minimizing the differences between the observed diffraction intensities and values computed from atomic positions in the model structure. For most protein crystals, the ratio of observable data to variable parameters is not very large, so that it is useful to introduce information about the local geometrical features that are constant among protein structures. This is achieved by the introduction of energy functions or geometrical restraints, and results in refined structures that have standard geometry in their detailed features.

Protein refinement is extremely laborious and computationally intensive. This stems from the relatively poor accuracy of initial models of protein structure and the poor convergence of the refinement methods, which tend to become trapped in local minima. As a result, few proteins can be refined automatically to a correct solution. Instead, parts that are modeled incorrectly have to be readjusted manually using interactive computer graphics and electron density maps whose coefficients represent the differences between the observed data and the density predicted by the current model. Initial model building with structure fragment libraries and molecular dynamics methods of x-ray refinement are two major developments that improve this situation. In the former approach, extended fragments derived from a library of highly refined protein structures are used to assemble the new structure, many features of which may not be apparent in detail from inspection of the initial electron density map. Conceptually, this is similar to the idea of incorporating standard group geometry into restrained refinement; and, indeed, models constructed with fragments refine rapidly, owing to the elimination of most modeling errors at the outset.

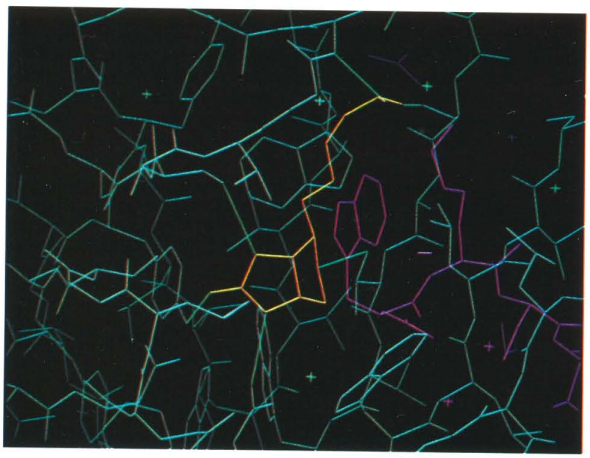
Molecular dynamics methods of protein refinement recently were described in CRAY CHANNELS.[2] The usual molecular dynamics simulations of proteins use a semiempirical potential function to describe pairwise atomic interactions and to define forces that govern the dynamics trajectory at normal temperatures. Molecular dynamics refinement basically involves the addition of a pseudopotential term, representing the difference between the Fourier coefficients observed from the crystal and those computed from the current model structure, to the usual interatomic potential. This, in effect, defines a trajectory for the system that eventually should arrive at the structure that best fits the data. The major advantage of the method is that it has a relatively large convergence radius, arising in part from its ability to break out of local minima when the simulations are run at high effective temperatures. Molecular dynamics refinement methods are very computationally intensive and have come into practical use only with the advent of supercomputers.

Despite the advantages of these advanced methods, neither is ultimately suited for automated protein refinement at the highest resolutions that most usefully reveal the details of inhibitor binding and protein-bound solvent structure. Thus, refinement remains a difficult procedure that must be performed independently for each receptor-drug complex.

Structural studies of drug binding

Determining how or where a structure will bind ligands is not normally possible from simple inspection. Moreover, even the application of the most sophisticated methods implicitly assumes that the binding site of the unliganded protein is unchanged when a substrate or inhibitor binds. In fact, this seldom seems to be the case. Fortunately, protein molecules generally retain full binding or catalytic activity in the crystalline state. This occurs because the molecules are interconnected only tenuously by weak ionic or hydrogen-bonded interactions, and otherwise are immersed in liquid water that fills continuous channels in the crystal lattice. Frequently, drugs or inhibitors simply can be diffused into the "native" protein crystal lattice and differences in x-ray scattering between the inhibited and native crystal directly transformed into a difference Fourier map that shows exactly how the drug is bound to the receptor molecule. Alternatively, it may be necessary to recrystallize the inhibitor-enzyme complex and solve its structure by molecular replacement methods.

The latter approach was used in a recent study of the tetrameric vitamin-binding protein streptavidin and its ligand, biotin.[3] This is an interesting system because the ligand binds with a dissociation constant of 10^{-14}M, making it among the strongest known protein-ligand interactions and of considerable utility in biological assay applications, where it is incorporated wherever a specific and irreversible link between biomolecules is required. From a comparison of the structures of streptavidin with and without biotin bound, it is apparent that the unusually high affinity of streptavidin reflects the participation of many factors that cooperate to allow formation of multiple hydrogen bonds between the protein and biotin. Important factors include the displacement of strongly bound water to



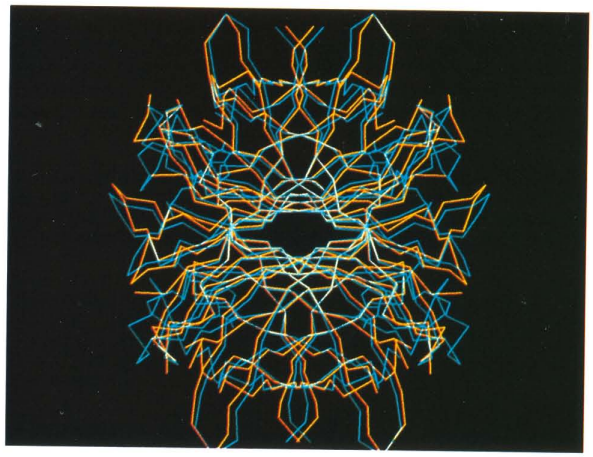


Figure 5. (Far left) The streptavidin binding site, showing the complexity of the interactions between biotin (in yellow) and surrounding residues.

Figure 6. (Left) A view showing a superposition of the streptavidin tetramer backbone with and without biotin bound.

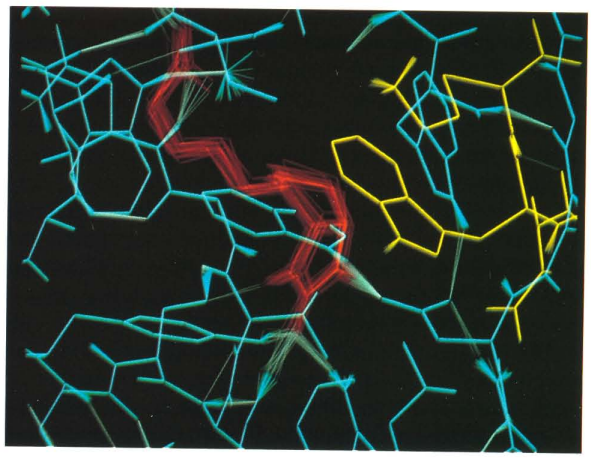


Figure 7. A simulation of biotin molecular dynamics in the surrounding environment of the streptavidin binding site.

reveal a polar binding site that binds to biotin ureido oxygen, whose oxyanionic form is, in turn, stabilized by an extended dipole array. These effects are enhanced by the ordering of a flap that sequesters the biotin from the surrounding aqueous environment (Figure 5). The whole binding process is accompanied by a substantial quaternary structural change in the tetrameric molecule (Figure 6). Measurements of the binding of biotin to streptavidin, or to a related protein avidin, suggest that the binding energy is essentially enthalpic and mainly involves the increased number of hydrogen bonds that the bound biotin makes owing to the special environment of the binding site. The important point is that the structure determination of the complex provides a detailed physical rationale of biotin binding affinity with features that can be tested explicitly by studying the binding and structures of analogs. At the same time, it is evident that virtually none of the important molecular effects or associated structural changes could be predicted accurately with available methods exclusively from a knowledge of the unliganded molecule.

Once a structural paradigm has been established for a class of inhibitors, a wide variety of modeling methods and data base search approaches can be applied to suggest rational modifications of the lead inhibitor or new compounds that might bind in analogous ways. Potential energy computations on the complex can provide detailed evaluations of the relative importance of specific interactions between the enzyme and inhibitor. In addition, free energy perturbation methods can provide estimates of binding free energies, although simpler simulations of complex molecular dynamics frequently can provide useful insight about ligand binding interactions (Figure 7).

Progress in structure-based design methods undoubtedly will accelerate as the accuracy and reliability of computational methods to assist rational design improve. In many cases, it seems clear that production of useful results will require increasing computational resources as required, for example, to simulate accurately the solvent environment around proteins. Nevertheless, it seems unlikely that purely computational methods will supplant crystallographic methods to determine protein ligand interactions in the near future. Instead, the most important computational advances in structure-based drug design are likely to be those that improve the speed and accuracy of protein structure determination.

About the authors

John J. Wendoloski is a member of the protein structure group at Du Pont, where he has been involved in applying ab initio methods and molecular modeling techniques to proteins and polymers. He received a B.S. degree in chemistry from the University of Scranton and a Ph.D. degree in chemistry from Yale University. Previously, he was on the staff of the National Resource for Computation in Chemistry.

F. Ray Salemme received a B.A. degree in molecular biophysics from Yale University in 1967. He earned a Ph.D. degree in chemistry from the University of California at San Diego in 1972 for x-ray structural studies of cytochrome c2. After 10 years at the University of Arizona in Tucson, Salemme joined Genex Corporation, where, as director of the Protein Engineering Division, he organized one of the first integrated research teams to engineer proteins using a combination of x-ray structural studies and computer-aided design methods. He joined the Du Pont Central Research Department in 1985, where he now heads a group whose objectives are the design and experimental development of engineered proteins and functional macromolecular assemblies.

References

1. de Vos, A. M., L. Tong, M. V. Milburn, P. M. Matias, J. Jancarik, S. Noguchi, S. Nishimura, K. Miura, E. Ohtsuka, and S. Kim, "Three-Dimensional Structure of an Oncogenic Protein: Catalytic Domain of Human c-H-ras p21," *Science,* Vol. 239, pp. 888-893, 1988.
2. Brünger, A. T., "Crystallographic Refinement by Simulated Annealing on Supercomputers," *CRAY CHANNELS,* Fall 1988.
3. Weber, P. C., D. H. Ohlendorf, J. J. Wendoloski, and F. R. Salemme, "Structural Origins of High-Affinity Biotin Binding to Streptavidin," *Science,* Vol. 243, pp. 85-88, 1989.

Temporal development of atherosclerosis in the human aorta

J. Fredrick Cornhill and Edward E. Herderick
The Ohio State University, Columbus, Ohio

Atherosclerosis is the most common disease in the Western industrialized nations. It causes most cardiovascular disease and results in more than 50 percent of all deaths in these nations. Atherosclerosis is the thickening of arterial walls in such a way that discrete lesions form on the insides of the arteries. This process gradually narrows the arterial lumen and finally occludes the arteries. The clinical manifestations of atherosclerosis are myocardial infarction (heart attack), cerebral vascular insufficiency and infarction (stroke), and peripheral vascular obstruction (gangrene).

Over the past several decades, researchers have learned much about the purported risk factors for heart disease, which include age, race, sex, lipoproteins such as LDL cholesterol and HDL cholesterol, cigarette smoking, hypertension (high blood pressure), diabetes mellitus, obesity, lack of physical activity, stress, and several genetic factors. Despite this gain in knowledge, little is known about the actual effects of these risk factors on the arteries themselves, particularly in the younger age groups. Virtually no quantitative data exist on the three-dimensional process of the development of atherosclerotic lesions — narrowing of the arteries — in humans. To develop such data, we have been analyzing aortal cross-sections using image analysis and computer equipment at The Ohio State University and the CRAY Y-MP computer system at the Ohio Supercomputer Center. This project constitutes a pilot study that is complementary to a larger cooperative project aimed at assessing the effects of various risk factors on the initiation and progression of atherosclerosis.

In 1985, a group of physicians and scientists in 14 laboratories organized a multicenter cooperative study to investigate the direct effects of risk factors on the arterial wall. The study, *Pathobiological Determinants of Atherosclerosis in Youth* (PDAY), is sponsored by the National Heart, Lung, and Blood Institute of the National Institutes of Health. The purpose of the project is to investigate the state of the arteries of persons between 15 and 34 years of age who have died from trauma and to correlate these observations with measured risk factor data.

As part of this study, the Laboratory of Vascular Diseases, of the Biomedical Engineering Center and the Department of Surgery at The Ohio State University, has been designated as the Morphometry Central Laboratory. The laboratory at Ohio State conducts quantitative morphometric studies on all macroscopic and microscopic images of arterial tissue entered into the study. The study encompasses about 100,000 images. These analyses utilize techniques of automated image processing wherever possible and result in the production of tabular and two-dimensional data. In a limited number of cases, three-dimensional data also are produced from magnetic resonance images (MRI). In addition, researchers have developed image analysis methods that allow gross images of individual arteries to be stretched to standard templates, using fiducial points located at anatomical landmarks, such as branches of different arteries.[1,2] This transformation to a standard arterial template allows the data obtained in a single case at a single point in time (a given age) to be grouped with other similar cases and then presented longitudinally. For example, a group of arteries taken from patients of different ages may be arranged in a sequence to give insight into the temporal development of atherosclerosis in the arterial system. The PDAY study has demonstrated that the creation of a longitudinal temporal progression from singular data obtained from cases of different ages greatly aids our understanding of the natural history of atherosclerosis.

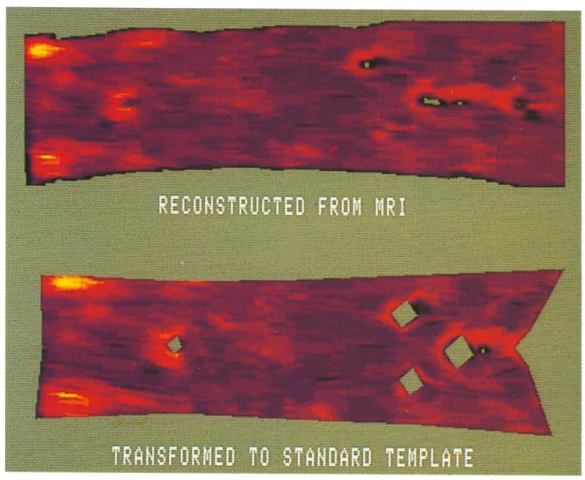


Figure 1. The upper image shows a two-dimensional reconstruction from MRI cross-sections of aortic wall thickness in the opened abdominal aorta of a 46-year-old male. The abdominal aorta was opened along the dorsal surface (top and bottom). The vessel is oriented such that the proximal end is to the right (celiac artery) and the distal end is to the left (aorto-iliac bifurcation). The thickness of the abdominal aorta is represented by the progressing color scale; dark purple: 0 to 1.5 mm, to yellow: 2.3 to 3.7 mm. The lower image shows the transformation of the thickness map in the upper image to the standard abdominal aortic template. The origins of the celiac (far right), superior mesenteric, paired right and left renal, and inferior mesenteric ostia (right to left) are presented as holes in the thickness map.

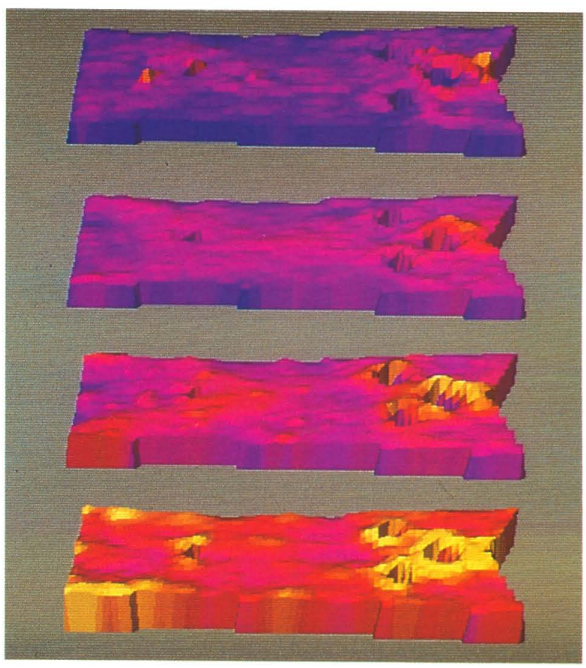


Figure 2. Three-dimensional representation, with thickness magnification, of the abdominal aorta from cases aged 5 to 9 years, 15 to 19 years, 30 to 34 years, and 45 to 49 years (top to bottom). Color scales of thicknesses are the same as in Figure 1.

In the pilot study complementary to the PDAY study, aortic wall thickness maps have been generated from MRI. MRI and image processing techniques have been developed to measure aortic wall thickness and thus identify regions that have increased significantly in thickness — the regions of the development of atherosclerotic lesion. Contiguous MRI cross sections of the abdominal aorta, obtained at autopsy, were taken at 3-mm intervals in the longitudinal direction using a 1.5-tesla magnetic resonance instrument in the Department of Radiology at The Ohio State University. The resultant voxel dimension is approximately 0.31 mm by 0.31 mm by 3.0 mm. The contiguous two-dimensional cross-sections were then reconstructed using an image analysis system comprising a Digital Equipment Corporation MicroVAX II and a Gould IP8400 image processor. Two-dimensional thickness maps of the opened abdominal aorta were produced, with pixel intensity representing aortic wall thickness at every location of the arterial lumen (Figure 1, top). These individual, reconstructed thickness maps then were transformed into a standard template (Figure 1, bottom). Images from several individuals then were grouped by age. The average thickness at every point on the aortic surface for each group was calculated.

This process yields one image for each age group in which the intensity of each pixel represents the average thickness at that position. Although these two-dimensional average thickness maps are useful in determining areas of interest, the data are by nature three-dimensional. A system has been developed, under a Cray Research grant, to visualize these data in three dimensions. The system consists of two programs. The first program takes the two-dimensional average thickness images and constructs a polygonal surface. The intensity of a pixel in the two-dimensional image determines the height of the surface and the color at that particular point. The second program renders the surface using a scan-line Z-buffer algorithm. This program features scaling of the axes, rotation of the surface, user-defined color maps, and various display models, such as wire frame, tiling, and Phong shading. The images shown were rendered on the CRAY Y-MP8/864 computer system at the Ohio Supercomputer Center and were displayed on the Gould IP8400 image processor at the Laboratory of Vascular Diseases.

The results of this pilot study for age groups 5 to 9 years ($n = 5$), 15 to 19 years ($n = 12$), 30 to 34 years ($n = 15$), and 45 to 49 years ($n = 14$) are presented in Figure 2 (top to bottom). This figure shows that the arterial wall increases in thickness, that is, atherosclerosis, with age, as indicated by both the three-dimensional representation and the change in color from dark purple to yellow. The figure also shows that some areas developed significantly more atherosclerosis than others. This is particularly true at the origin of the celiac, superior mesenteric, and right and left renal arteries (right) and on the dorso-lateral surface of the distal abdominal artery just proximal to the aorto-iliac bifurcation (left). Figures 3 and 4 demonstrate the ability of the program to display the surface from various angles.

The display of static images, shown in Figure 2, is appropriate for binary groups, such as smokers versus nonsmokers. However, with continuous risk factor variable data, such as LDL cholesterol and age, the most effective method for displaying such data is to generate an animated sequence. The Ohio Visualization Laboratory at the Ohio Supercomputer Center provides the equipment necessary to transfer sequences of images to video tape. The development of animated sequences that show the effects of the alteration of continuous variables on the development of atherosclerosis holds great promise for our understanding of the dynamic role of risk factors in atherogenesis. Such studies are underway in cooperation with the Ohio Supercomputer Center.

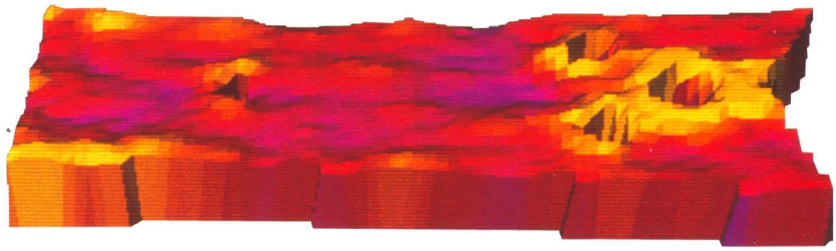
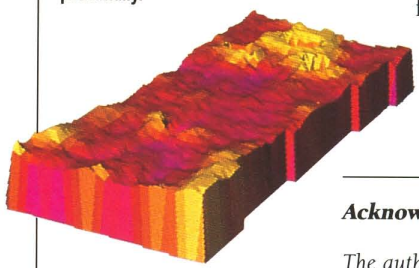


Figure 3. Three-dimensional representation of the abdominal aorta of the 45 to 49 year age group as viewed from the dorsal surface.

Figure 4. Three-dimensional representation of the abdominal aorta of the 45 to 49 year age group as viewed from the aorto-iliac bifurcation, looking proximally.



We now are able to generate video sequences in which we can "watch" the initiation and development of atherosclerosis and observe the effects of risk factors in an entirely new way. The ability to generate sequences containing hundreds of images in a timely fashion gives us a powerful tool in the study of atherosclerosis. And such work can be accomplished only through the use of computer systems such as the CRAY Y-MP system. ■

Acknowledgments

The authors thank Suzanne Smith for producing the original version of the software and J. Christopher Forman for modifying it for use on the CRAY Y-MP system. The work described in this article was supported in part by grants from the Public Health Service (HL 33760 and RR 04115), Cray Research, Inc., and the Office of the Vice President for Research and Graduate Studies of The Ohio State University.

About the Authors

J. Fredrick Cornhill is director of the Biomedical Engineering Center and is professor of surgery and pathology at The Ohio State University. He received a B.Sc. degree in biophysics from the University of Western Ontario in 1972 and a D. Phil. degree in engineering science from the University of Oxford in 1975. He has been on the faculty of The Ohio State University since 1976 and currently directs the Laboratory of Vascular Diseases and the Laboratory of Vascular Biology.

Edward E. Herderick is a systems programmer in the Laboratory of Vascular Diseases at The Ohio State University. He received a B.S. degree in statistics from The Ohio State University in 1978. In 1980 he joined the Laboratory of Vascular Diseases, where he currently directs the development of image processing software.

References

1. Cornhill, J. F., W. A. Barrett, E. E. Herderick, R. W. Mahley, and D. L. Fry, "Topographic Study of Sudanophilic Lesions in Cholesterol-Fed Minipigs by Image Analysis," *Arteriosclerosis*, Vol. 5, pp. 415-426, 1985.
2. Cornhill, J. F., E. E. Herderick, and H. D. Starey, "Topography of Human Aortic Sudanophilic Lesions," *Monographs on Atherosclerosis*, (in press).

Pseudoreceptor modeling
An experiment in large-scale computing

Shashidhar N. Rao and James P. Snyder, Searle Research and Development, Skokie, Illinois

Broad-based computational groups are responsible for a diverse set of applications and development tasks. The workers in such a group employ a supercomputer as a partner in a network of other platforms such as superminicomputers, stand-alone workstations, parallel platforms, personal computers, and most often a VAXcluster. Generally, only the most numerically intensive computations are presented to the largest device, where many other users are competing for queue position. The Drug Design Group at Searle is using a CRAY X-MP/116se supercomputer to explore approaches to compound design not possible on smaller computers. One such area of study concerns the interaction of drug molecules with proteins ordinarily embedded in a membrane bilayer that constitutes the surface of cells.

Drug action at the level of cell function and membrane transport is mediated by a series of microscopic events. The essential trigger in the multistage process generally is believed to be a highly specific recognition of the drug molecule by a macromolecular species. In the best-behaved situations, receptor binding characterized by an affinity constant, K_D, is accompanied by a functional correlate such as nerve-cell firing, muscle-strip contraction, or the production of second messengers inside the cell. Drug design, when it integrates computational methodology, most often focuses on the binding event and manipulates structure-activity data in three basic ways.

Modeling the binding of drugs to biological receptors

Possession of a three-dimensional structure or model of the therapeutically relevant macromolecule — an enzyme, DNA fragment, or glycoprotein — is a circumstance rich with information, especially when the biopolymer is embellished with a tight-binding ligand. Active site-docking and a range of molecular mechanics and dynamics experiments on complexes containing both actual and putative drug molecules are possible. This procedure is called *receptor fitting*, a method in which a small molecule is molded to the requirements of the macromolecule's binding locus. At the other extreme, quantitative structure activity relationship methods, which rely on linear regression or principal component analysis, attempt to summarize bioactivity (for example, K_D's) compactly in the form of multiterm equations. Past work has relied very little on the use of whole-molecule structural features. Physiochemical parameters and substituent constants have been used most frequently to obtain correlations. However, structure-based descriptors and volume and electrostatic-potential components are becoming the preferred biocorrelates. Notable in this respect are the recently reported comparitive molecular field analysis (CoMFA) and hypothetical active site lattice (HASL) approaches.

Intermediate in the attempt to predict biological activity from structure is a family of methods called *receptor mapping*. Molecular properties of the receptor are projected outward in a complementary fashion from the characteristics of a series of active ligands to produce a pharmacophore, which is a three-dimensional arrangement of functional groups common to a bioactive set of compounds whose orientation at the active site is essential for recognition and binding. The "active analog" approach,[1] distance geometry applications,[2,3] and APOLLO[4] are representative. All of these schemes rely on a comparison of molecules active at a given receptor and seek a bioactive conformational motif as the predictive element of the pharmacophore. In some cases, molecular fragments are positioned in three dimensions about the construct to represent side-chain site binders attached to the receptors. With few exceptions, these moieties are

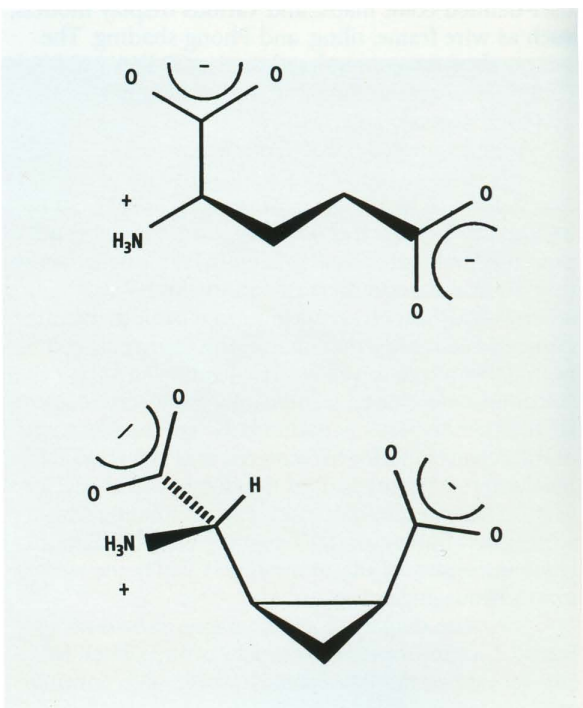


Figure 1. Schematic drawing of the ligands bound to the pseudoreceptor: (top) glutamic acid, (bottom) cyclopropanated glutamic acid.

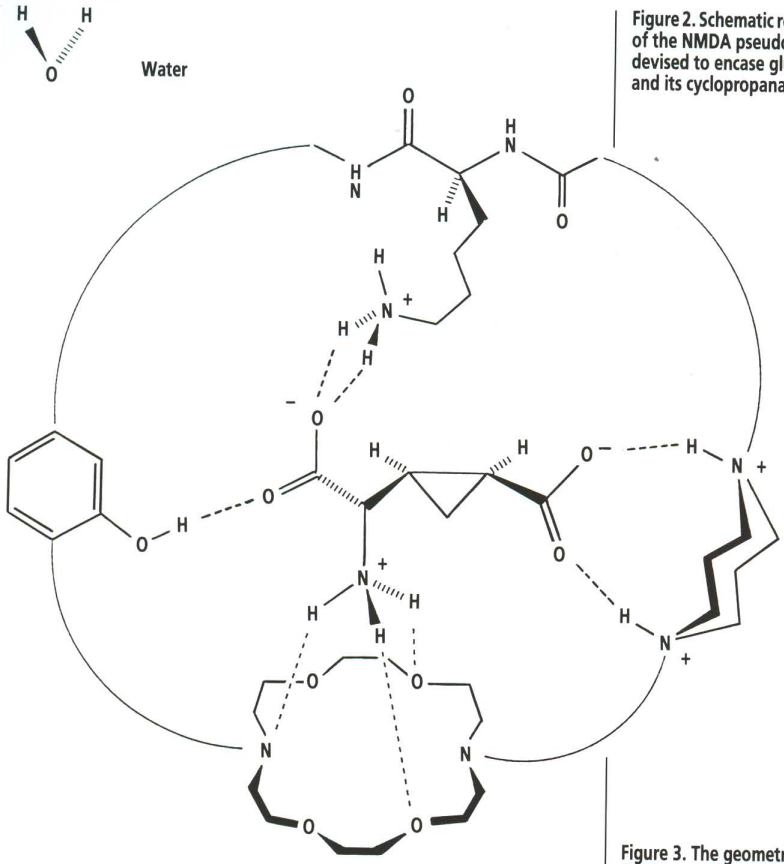


Figure 2. Schematic representation of the NMDA pseudoreceptor devised to encase glutamic acid and its cyclopropanated analogs.

Figure 3. The geometry-optimized NMDA pseudoreceptor fitted with SRS-cyclopropyl-glutamic acid; hydrogen bonds displayed.

adornment, illustrative only of the spatial constitution of the pharmacophore.

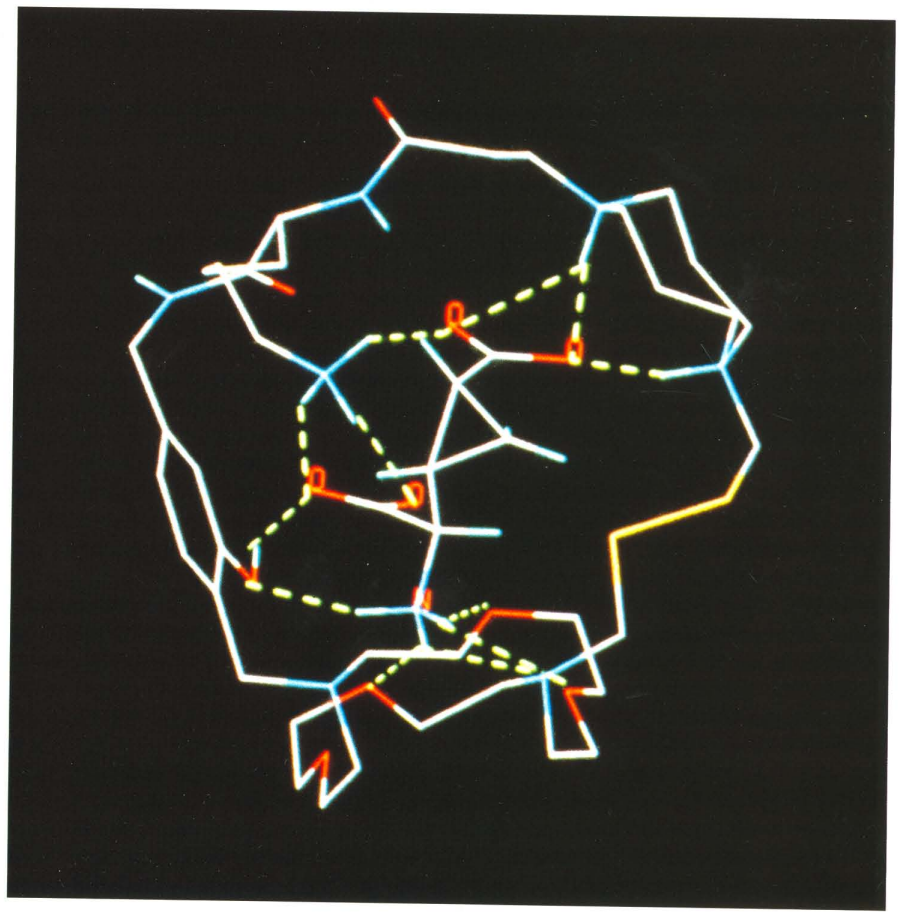
The pseudoreceptor bridge between fitting and mapping

Missing from the tool kit of the drug designer is a receptor mapping procedure cast in the framework of receptor fitting. One strategy is to design a structurally valid, but highly truncated active-site pocket around a potent receptor ligand. Once such a complex has been refined, replacement of the original substrate with suitable lead candidates can provide novel complexes testable for steric and electrostatic fit. Subjecting the pseudoreceptor couple to methods capable of yielding relative free energy differences ($\Delta\Delta G$) raises the possibility for predicting relative binding constants ($\Delta K_D(rel)$) quantitatively. Such an approach is particularly critical for the growing list of receptor-ligand interactions driven by entropy.

We have initiated some numerical experiments along these lines for analogs of the ubiquitous central nervous system activator glutamic acid (Figure 1, top). The three major pharmacological receptors to which it is known to bind are NMDA (N-methyl-D-aspartate), Quis (quisqualate), and Kainate. While the latter are primarily research entities at present, the NMDA site participates in a matrix of regulatory actions in concert with the phencyclidine (PCP) and glycine receptors coupled to a set of ion channels. The entire complex has been implicated in neurological disorders such as memory loss, schizophrenia, anxiety, stroke, and cerebral ischemia. The discovery of potent and selective NMDA ligands is considered to be one approach to treating these disorders. Like the overwhelming majority of membrane-bound receptors, the structure of neither the full protein complex nor the NMDA component is known.

Charged neurotransmitters as probes of computational accuracy

One diastereomer of cyclopropanated glutamic acid (SRS-CPG) (Figure 1, bottom), has been shown to bind to the NMDA receptor about 300 times more tightly than natural S-glutamic acid.[5] This corresponds to a relative free energy gain of 1.3 Kcal/mole at 2°C. Apart from their biological significance, the glutamate analogs are a superb class of structures for evaluating the pseudoreceptor complex. They are small molecules with complex but manageable conformational surfaces. Each carries three full charges, preventing a straightforward treatment by traditional gas phase methods. Finally, the synthetic derivative shown in Figure 1 (bottom), is semirigid, providing some clues to the shape of the NMDA receptor agonist binding cleft. For these reasons, we decided to use it as a ligand template for pseudoreceptor design with the constraint that all three charged groups are satisfied by strong salt-bridge formation. The study was initiated with a conformational analysis of SRS-CPG to identify rotamers consistent with an analogous family of low-energy conformations of naturally occurring S-glutamic acid (S-Glu). Figures 2 and 3 illustrate the receptor tailored to the selected CPG conformation. The tightness



of the fit is depicted by surrounding the ligand with an expanded molecular surface (Figure 4).

The $NH3^+$ group is complexed to an 18-crown-6 analog at its base. The negative charge of one COO^- is satisfied by a lysine side chain and a phenol-OH. The second COO^- is covered by a di-NH^+ eight-membered ring guanidine mimic. Each of these receptor fragments is linked by various molecular spacers. Once the complex was fabricated, it was geometry optimized in the gas phase with the MM2 force field. Subsequently, glutamate enantiomers (S and R) were placed in the optimized receptor in low-energy conformations accommodating the binding site charge distribution (Figure 5). Again, MM2 optimization was performed. The triple of gas phase complexes served as input for the relative free-energy calculations.

Each of the MM2 optimized ligands and complexes was reoptimized using the AMBER-UCSF force field, first in the gas phase with a distance dependent dielectric, then within a solvent shell consisting of over 400 explicit Monte Carlo waters (Figure 6). The resulting energies are not free energies and therefore cannot be directly compared for different chemical constitutions. Since the ultimate goal of the study is to rationalize the experimental K_D's in terms of theoretical models, the molecular mechanics calculations need to be supplemented with molecular dynamics coupled with free energy calculations. The latter, of course, yield both enthalpic (ΔH) and entropic (ΔS) contributions.

Free energy perturbation and relative receptor binding potencies

The molecular-dynamics-based free energy perturbation treatment (FEP) has proved to be a quantitative tool for predicting differences in ΔG between closely related ligands for certain enzymes.[6] Figure 7 shows the central problem in the evaluation of the differential free energy of the binding of two closely related ligands A and A′ to a common receptor. Experimental studies measure differences in free energies represented by the horizontal arrows. Such processes are difficult to simulate theoretically. On the other hand, modeling the "processes" represented by vertical arrows is relatively straightforward. In this way the calculated free energies can be related to experimentally observed energy differences. The left-hand side of the thermodynamic cycle accounts for the differential solvation of the two ligands; the right-hand side to relative complex formation potentials.

Theoretical simulation of the hypothetical processes represented by vertical arrows presents the following challenges. First, the force fields employed treat electrostatic interactions simplistically by considering only charge-charge interaction terms with a constant dielectric of 1.0. In such a scheme, electrostatic effects usually are overemphasized, particularly for systems containing a number of charged groups. Explicit solvation helps to mediate the strong Coulombic interactions. For this reason, both complexes and the ligands have been solvated by a sphere of over 400 water molecules.

Second, the protocol for evaluating free energies involved molecular dynamics simulations

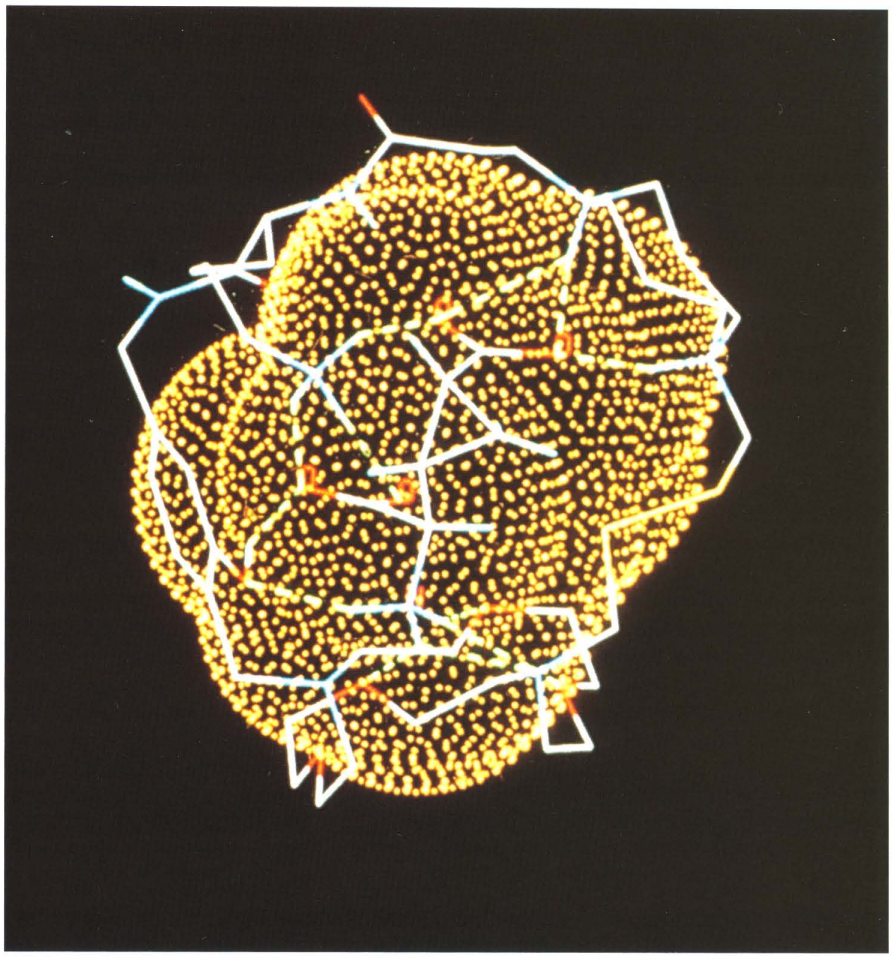


Figure 4. The NMDA pseudo-receptor and the extended-radius surface of the SRS-cyclopropyl-glutamic acid indicating the tight fit of receptor and ligand.

over a period of 20 to 80 psecs depending on the structures being evaluated. For example, calculations on the relative free energy differences of cyclopropanated glutamic acids were initiated with 10 psecs of molecular dynamics equilibration. The perturbations then were carried out using the slow-growth technique for a period of 60 psecs in each direction of the perturbation process. Molecular dynamics and free energy calculations over such periods with more than 1300 atoms, as in our complexes, are computationally intensive tasks. The computational intensity is further highlighted by the fact that very short time steps of

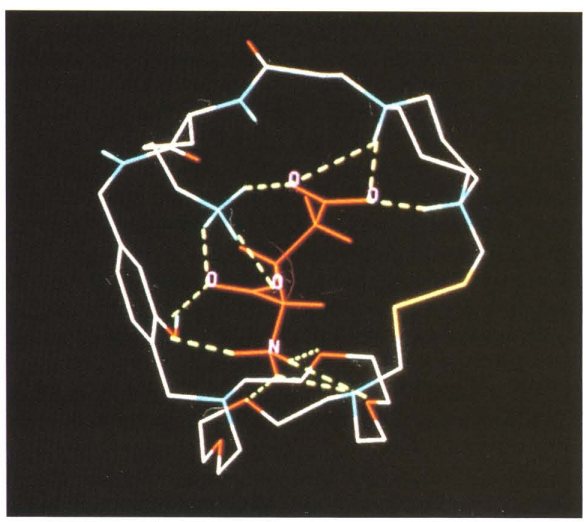


Figure 5. A low-energy conformation of S-glutamic acid in the optimized receptor-ligand complex; hydrogen bonding is identical to that observed for the more potent SRS-CPG agonist.

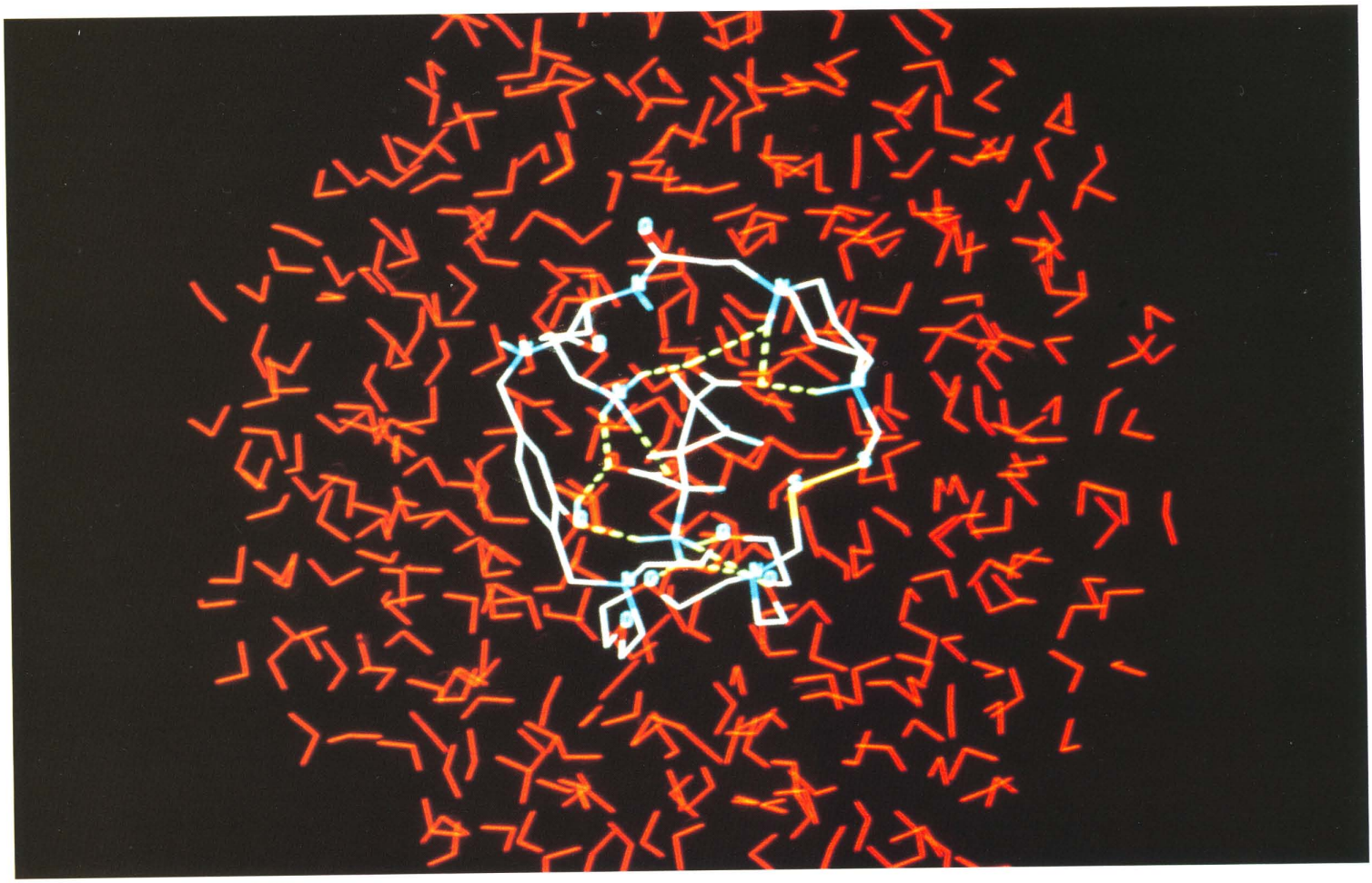


Figure 6. The NMDA pseudo-receptor and ligand in a sphere of waters during free energy perturbation.

0.0005 psecs and a corresponding large number of molecular dynamics steps were traversed in the simulations.

The problem of conformational variation

In all the previous applications of the free energy perturbation methods for evaluating binding energies of ligands to receptors (macromolecules), models for the two components were obtained from x-ray crystallography. Such models have an inherent advantage in that a large fragment of the macromolecule component typically can be frozen during the simulations, ensuring that conformational flexibility is restricted — at least in the receptor. In a few cases of enzyme-substrate binding, it has been demonstrated that even the ligand undergoes only minimal conformational changes during the course of perturbation. In the present study, although we have no x-ray structure for the NMDA receptor, the computer graphics-based hypothetical construct lends itself directly to the free energy analyses. No constraints were placed on the movements of receptor atoms.

To compare compounds 1 and 2 with variations in both conformation and atom count, we have adopted a ligand "disappearance" strategy for the right and left legs of the thermodynamic cycle (Figure 7). Two aspects are involved: the charges are changed gradually during the course of the perturbation from the quantum mechanically derived values to 0.0 on each of the atoms, and the van der Waals parameter ε is altered from its standard force field value to 0.0 for each of the atoms in the system. Thus, the ligand molecule under study is slowly deprived of its steric and electrostatic properties while the micro-environment constantly reoptimizes its structure to fill the resulting "gaps." The process is reversed in a subsequent calculation with reappearance of the ligand and concomitant reorganization of the surrounding milieu.

In most earlier free-energy studies of ligand-receptor binding, the approach of bidirectional calculations has been employed. Uncertainties in the free-energy differential characterizing the end-point thermodynamic states due to conformational changes are thereby minimized. Typically, a portion of the receptor is held fixed. By contrast, in the present study, in which the entire model receptor is allowed to move, we observe nontrivial conformational changes about

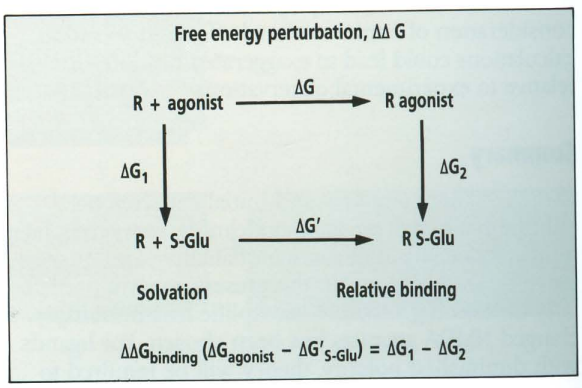


Figure 7. Thermodynamic cycle relating the measured free energies of binding to a common receptor (ΔG1 and ΔG2) for two structurally similar ligands A and A'.

some of the receptor bonds. Similar receptor reorganization occurred in earlier studies at the active site of an enzyme.[7] In spite of the torsional flexibility, the conformational changes in the NMDA pseudoreceptors are similar for various ligands. Furthermore, the solvated ligand and complexes experience free energy variations within ± 2 Kcal/mole for the forward and reverse runs.

The work was initiated on a CRAY X-MP/48 system running under the UNICOS operating system. A typical set of calculations performed to evaluate free energies consumed about 18 to 20 hours for the combined forward and reverse directions on the Cray system. The Cray system's performance was 60 times faster than that of a VAX/8650 computer under the VMS operating system when applied to an identical set. The calculations described here and in an accompanying broader study with a CRAY X-MP/116se system would not have been contemplated without the capacity of the supercomputer.

The three ligands adopt an extensive ensemble of conformations in the experimental aqueous milieu. The present work has adopted an Occam's razor strategy of carrying out "solution" calculations on a representative conformation of the three zwitterions; namely, that which arises from the AMBER optimized ligand-receptor complex embedded in the 400-molecule water bath. Results are shown in Table 1.

The SRS-CPG compound that binds to the NMDA receptor better than S-Glu is predicted to do so within 0.8 Kcal/mole; slightly outside the experimental error of 0.3 Kcal/mole. R-Glu, on the other hand, binds less efficiently to NMDA, and the calculations mimic this qualitatively. However, the relative difference is underestimated by an order of magnitude. This can be attributed in part to the fact that the pseudoreceptor was tailored to SRS-CPG 2 with excessively stringent torsional constraints (Figures 2-4). Provision of a more flexible pseudopocket is expected to preserve the features of the very strong binding ligands but to accommodate the poorly binding ones more effectively.

The calculations to date have not taken into account the conformational flexibility of the uncomplexed ligands. Even though the cyclopropanated glutamates are conformationally restricted relative to glutamic acid, they exhibit flexibility about the C^{α}-C^{β} bond. We are investigating the effects of different conformations on solvation energies for some of the isomers of CPG. Preliminary results indicate that some conformations are solvated preferentially and more strongly than others. It is therefore possible that, with oversimplifications in the electrostatic treatment in the methods employed, the lack of explicit consideration of conformational effects in solvation calculations could lead to exaggerated binding energies relative to experimental observations.

Summary

These tentative and initial experiments indicate that pseudoreceptor modeling by supercomputer ultimately may lead to routine qualitative and quantitative K_D comparisons. In the present study a particularly demanding situation exemplified with multiply charged NMDA agonists has been chosen. For ligands with diminished polarity, theory will be required to quantify a range of possible hydrophobic effects in this context, as well. Before the method can be considered an applications tool, several learning curves must be surmounted. Larger data sets of ligands must be treated, and a general protocol for constructing pseudoreceptors is needed. The question of receptor flexibility requires evaluation, as does the importance of the conformational profiles of ligands in "solution." Further, method development around intermolecular force field interactions, in particular the introduction of polarization terms, may be required. Supercomputers constitute one of the platforms that satisfy the enormous CPU demands of the FEP treatment and thereby help to translate computer experiment into routine practice.

Compound	ΔG3	ΔG4	ΔΔG(calc)	ΔΔG(exp)[a]	Ki[b]
S-Glu	0.0	0.0	0.0	0.0	0.62
R-Glu	0.2	17.1	–16.9	–1.7	14.5
SRS-CPG	0.3	–1.7	2.0	1.3	0.059

(a) Calculated from $\Delta G = -RT\ln(Ki)$; $\Delta\Delta G = (\Delta G(\text{agonist}) - \Delta G(S - glu))$

(b) Binding assays were performed by displacement of L – ^{3}H – glutamic acid at °2 C.

Table 1. Comparison of FEP calculated and experimental relative free energy differences for R and S glutamic acid and SRS-cyclopropanated glutamic acid following the thermodynamic cycle of Figure 2; Kcal/mole.

Acknowledgments

The authors thank Cray Research for a generous amount of computer time on a CRAY X-MP/48 system in the early stages of this work.

About the authors

Shashidhar N. Rao joined Searle in 1987 after a postdoctoral research fellowship at the University of California, San Fransisco. He earned a Ph.D. degree in molecular biophysics from the Indian Institute of Science, Bangalore, India, in 1983.

James P. Snyder joined Searle as head of the Drug Design Group in 1984, following three years as a senior research fellow at Merck Sharp and Dohme. He previously held teaching positions at the University of Copenhagen, Denmark, and at the Belfer Graduate School of Science, Yeshiva University, New York. He was awarded a Ph.D. degree in organic chemistry from Cornell University in 1965.

References

1. Dammkoehler, R. A., S. F. Karasek, E. F. Berkley Shands, and G. R. Marshall, *The Journal of Computer-aided Molecular Design,* Volume 3, pp. 3-21, 1989.
2. Ghose, A. K., G. M. Crippen, G. R. Revankar, P. A. McKernan, D. F. Smee, and R. K. Robins, *The Journal of Medical Chemistry,* Volume 32, pp. 746-756, 1989.
3. Sheridan, R., R. Nilakantan, J. S. Dixon, and R. Venkataraghavan, *The Journal of Medical Chemistry,* Volume 29, pp. 899-906, 1986.
4. Koehler, K. F., D. P. Spangler, and J. P. Snyder, "Pharmacophore Identification Through Molecular Similarity," poster presentation at the 196th National Meeting of the American Chemical Society, Los Angeles, California, Sept. 25-30, 1988.
5. Pellicciari, R., B. Natalina, M. Marinozzi, L. Selvi, C. Chiorri, J. Monahan, T. Lanthorn, and J. P. Snyder, *Neurology and Neurobiology,* Volume 46, pp. 67-70, 1988.
6. Singh, U.C., F. K. Brown, P. A. Bash, and P. A. Kollman, *The Journal of the American Chemical Society,* Volume 109, pp. 1607-1614, 1987.
7. Bash, P. A., U. C. Singh, F. K. Brown, R. Langridge, and P. A. Kollman, *Science,* Volume 235, pp. 574-576, 1987.

Protein structure determination from magnetic resonance data

David A. Case, Research Institute of Scripps Clinic
La Jolla, California

During the past decade, researchers have made remarkable progress in applying nuclear magnetic resonance spectroscopy (NMR) to the study of proteins.[1] NMR is applied to the determination of protein structures by placing a sample of protein in a strong magnetic field and recording the absorption of radiation in the radio-frequency range. The resulting spectrum can reveal a variety of nuclei; the most common experiment is sensitive to protons (hydrogen nuclei) in the sample. The introduction of very strong magnets (about 11 tesla), along with novel excitation and recording techniques, enable researchers to resolve signals from individual protons and to make assignments of each of the thousands of peaks in a spectrum to particular chemical locations in a protein. The assignments per se are of limited interest, but they enable the interpretation of "cross-peaks" in the spectrum that provide information about the distances between pairs of proteins. Using a CRAY X-MP EA/116 computer system, scientists at the Research Institute of Scripps Clinic are determining three-dimensional structures of proteins from such NMR information.

The basic approach

Figure 1 depicts the flow of information for determining solution structures from NMR data. The methods by which individual peaks in the spectrum are assigned to sequence-specific locations in the chemical structure are outside the scope of this article. Such assignments allow one to interpret nuclear Overhauser effect (NOE) spectra. In this experiment, protons that are close together (less than about 5 angstroms) show a cross peak in a two-dimensional spectrum.

As indicated in Figure 1, the identification of a small number of such close distances often allows the general pattern of secondary structure (for example, of helices, sheets, and turns) to be assigned, because different patterns of secondary structure have characteristic sets of short distances. A more quantitative interpretation of the NOE information requires one to consider hundreds to thousands of cross peaks — a task for large-scale optimization via supercomputer.

This procedure begins by interpreting cross peaks in terms of fairly qualitative distance constraints between protons: a "strong" cross peak may indicate that two protons are within 2.5 angstroms of each other, while a "weak" peak will be assigned an upper bound of 5 angstroms. Many other distance constraints can be inferred from the known chemical structure or from geometrical considerations such as the triangle inequality (which states that no side of a triangle can be longer than the sum of the other two sides.) One thus obtains up to thousands of distance constraints, and the task is to determine which protein conformations are consistent with these constraints. Invariably, this problem is under-determined, so a "family" of conformations that satisfies the experimental constraints exists. Typically, these families are depicted graphically by a superimposition of a number of structures that satisfy the constraints and that approximately represent the complete distribution of allowed conformations.

Distance geometry and simulated annealing

The most common mathematical procedure by which distance constraints are converted into three-dimensional structures is called "distance geometry," a method that involves diagonalization of the metric matrix, which can be constructed solely from distances.[2] Geometrically, this procedure projects the structure

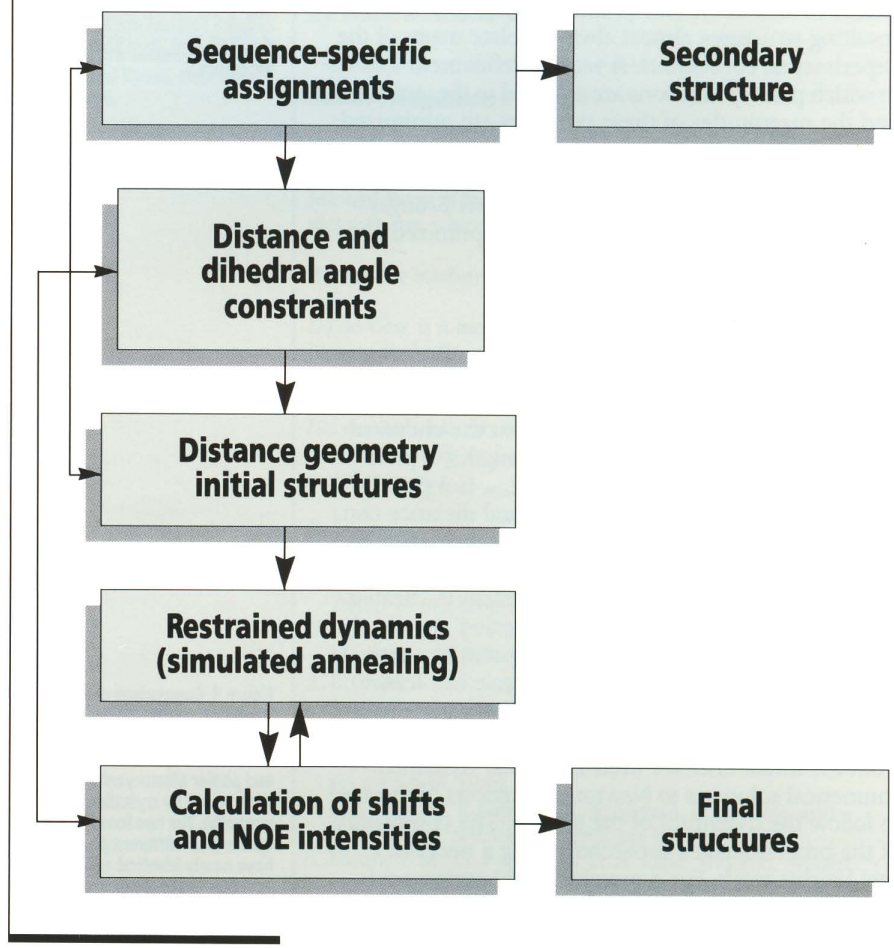


Figure 1. Flow of information in an NMR structure refinement.

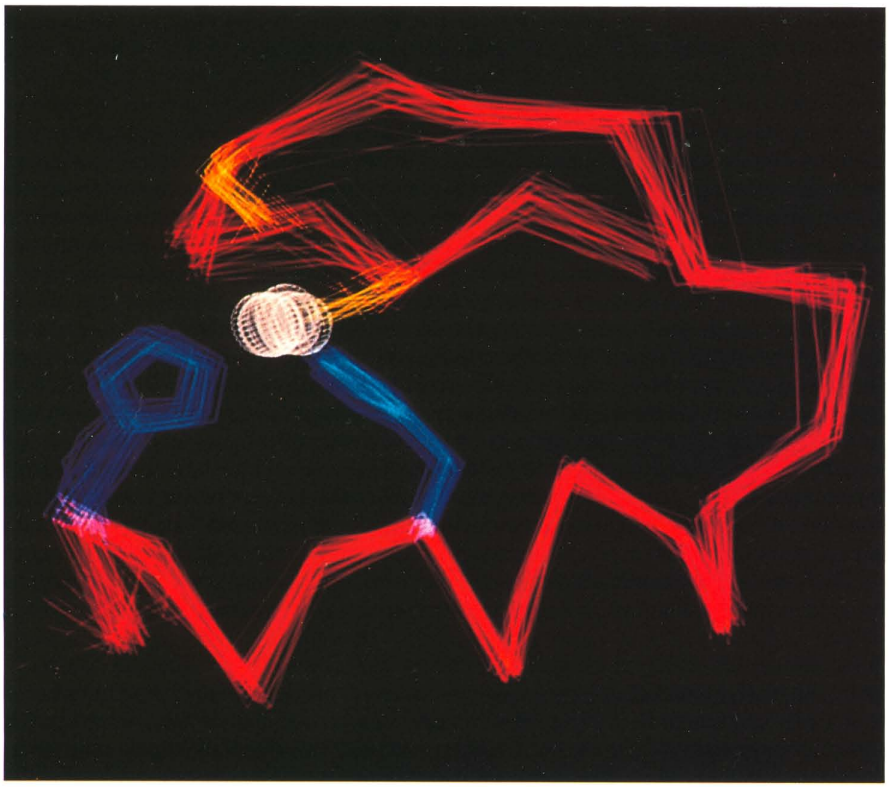


Figure 2. Superposition of 37 structures of a zinc finger domain. A view of the cα chain tracing is shown, along with the zinc atom (white) and its cysteine (yellow) and histidine (blue) ligands. Data from Reference 5.

from a high-dimensional space (in which arbitrary distance relations can be accommodated) into ordinary three-dimensional space. This projection procedure entails considerable loss of information, and the resulting structures almost always violate many of the experimental constraints. A second refinement follows in which penalty functions are assigned to the violations, and the magnitudes of these penalties are minimized through adjustments to the protein structure. These adjustments are carried out through a simulated annealing technique similar to that used by Axel Brünger of Yale University.[3] The function to be optimized is represented as

$$E(x) = E_{MM}(x) + E_{NOE}(x),$$

where x represents the coordinates of the atoms in the protein. E_{MM} is a "molecular mechanics" potential that incorporates what we know about the chemical structure of proteins, such as bond lengths, typical angles, and nonbonded exclusions. E_{NOE} is a penalty function that is zero if the experimental distance constraints are satisfied, and that becomes increasingly positive (unfavorable) as the constraints are violated. With this combined potential, the system is "heated" to a high temperature and then "annealed" by slowly removing thermal energy from the system. As with analogous physical systems, such a cycle can lead to lower energy conformations, even when large energy barriers must be crossed to reach the final configuration from the initial one. We used molecular dynamics (numerical solutions to Newton's equations of motion) to follow the dynamics of the system. The complexity of the process makes supercomputing a necessary tool for this work. For a protein of 100 amino acids, each heating and cooling cycle may take about 30 minutes of time on a CRAY X-MP system processor, and this procedure has to be repeated enough times to allow one to map a reasonable description of the "family" of allowed structures.

Even more computer power may be needed for the next stages of refinement (Figure 1, bottom left), in which detailed comparisons of calculated and experimental spectra are used to check the validity of the resulting structures and to refine them further. The essential complication in this step is that the magnitude of a particular cross-peak depends not only upon the distance between the two protons it connects (as assumed above), but also upon the locations of all other protons in the molecule. This dependence upon other protons results from spin-diffusion, a process by which the magnetic orientation of a particular spin diffuses throughout the molecule. We recently have developed a method to include spin diffusion effects into a simulated annealing calculation, and we are testing it on small proteins and oligonucleotides.[4] The method requires diagonalization of a rate matrix that describes the spin diffusion process and manipulation of the resulting eigenvectors. Hence it also is well-suited to vector architectures.

Unfortunately, the difficulty of the problem grows as the fourth power of the number of protons in the molecule, so that the attempt to solve such problems can outstrip even supercomputer capabilities. With the current implementation we can handle systems with about 300 protons and are working on approximation schemes that will enable us to tackle larger systems.

Examples

The general status of protein structure determination by NMR recently has been reviewed.[1] Several dozen protein structures have been determined by NMR methods, and the rate at which new structures are being determined is accelerating. Figures 2-4 show

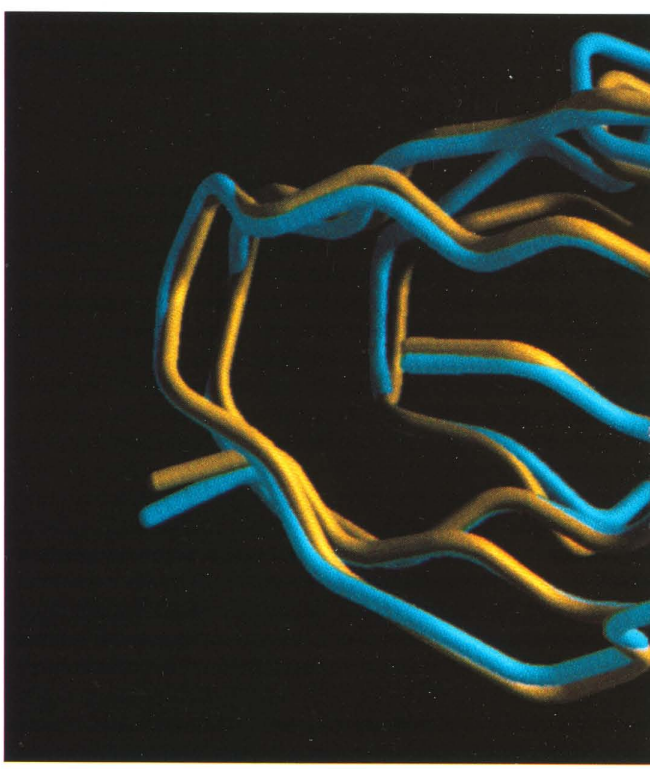


Figure 3. Comparison of the backbone configurations of French bean plastocyanin (yellow, determined by NMR methods) and poplar plastocyanin (blue, determined by crystallographic methods). The two forms of plastocyanin from different plants have nearly identical amino acid sequences. Data from Reference 6.

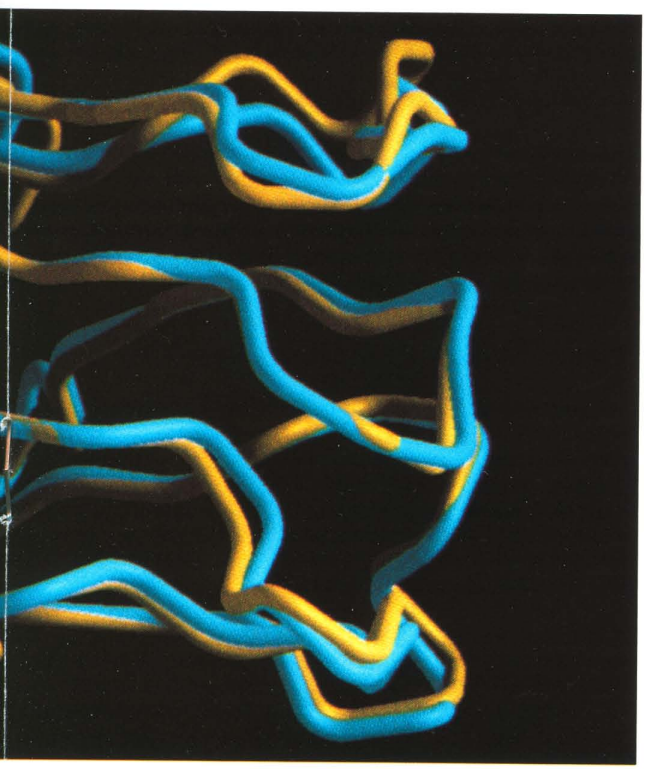
examples of proteins recently solved by this procedure on the Cray computer system at Scripps Clinic.

Figure 2 shows a superposition of 37 independently derived structures for a "zinc-finger" domain protein.[5] Zinc fingers are regulatory proteins that control transcription and other aspects of DNA processing. The remarkable congruence of all 37 structures shows that the resolution that can be obtained by NMR methods can begin to rival that of x-ray crystallography. For the zinc finger, the uncertainty in position for atoms in the polypeptide backbone is about 0.5 angstroms.

Figure 3 shows results for a much larger protein, plastocyanin, which has 99 amino acids.[6] Plastocyanins are copper-binding proteins that serve as electron transport agents in photosynthesis. The figure compares the NMR-determined structure with that obtained on a nearly identical protein by x-ray crystallography. It is clear that the two independent methods are achieving nearly identical results, and that the structure seen in the crystal is indeed relevant for discussions of protein function in solution. Figure 4 demonstrates the superposition of a few structures near the copper-binding site, showing the extent to which even side chains can be confidently located if enough NMR information is available.

Conclusions

The structural studies described here would not have been possible a few years ago, primarily because of limitations in NMR methodology, but also because of the magnitude of the computations involved. As the precision of the experimental data has improved, the need has grown for more sophisticated (and time-consuming) refinement techniques. Some of the techniques outlined here are still too computer-intensive to be applied on a routine basis, and Scripps Clinic researchers are working to develop less costly procedures.



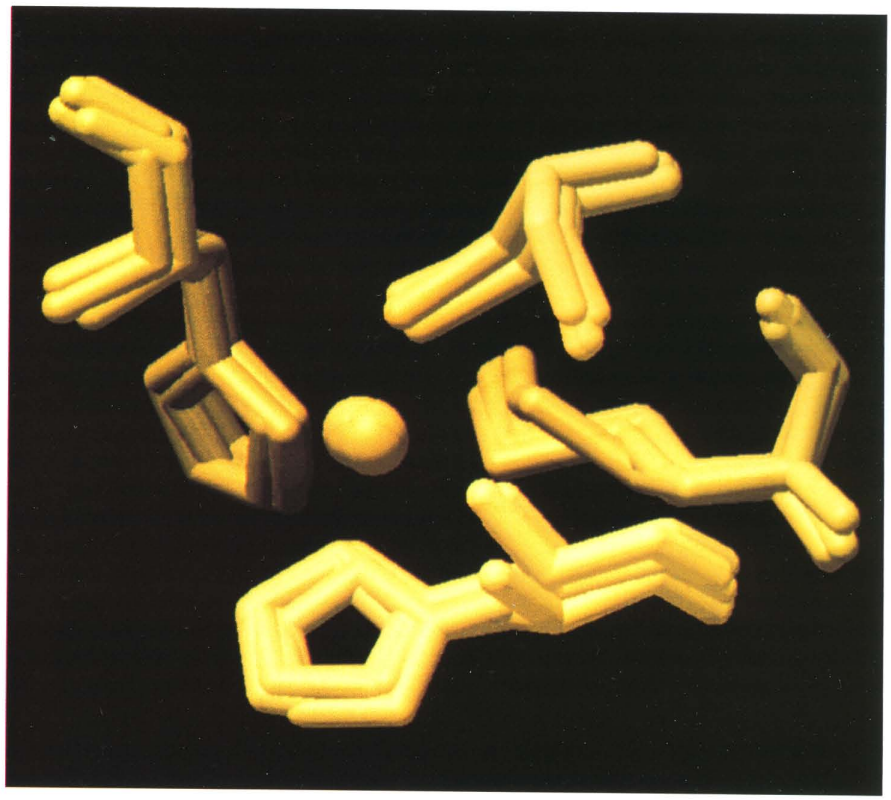


Figure 4. Superposition of three NMR-determined structures near the copper-binding site of plastocyanin. The copper atom is represented by the sphere in the middle, which is bound to two histidine, one methionine, and one cysteine ligand. Data from Reference 6.

These efforts, plus a continuation of the drop in the effective cost of computation, should allow continued progress in the determination of biopolymer structures by NMR methods. ▬

Acknowledgments

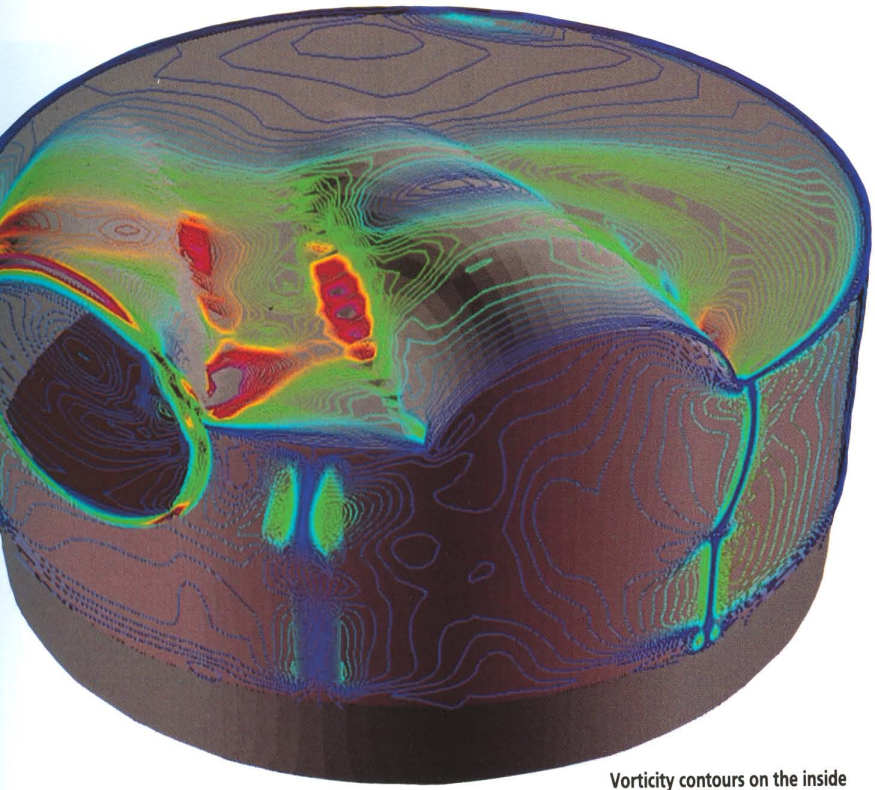
The author thanks Garry Gippert, Ping Yip, and Peter Wright for many helpful discussions. Figures 3 and 4 were prepared by Michael Pique of the Research Institute of Scripps Clinic, using mcs software, which was written by Mike Connolly.

About the author

David Case is a member of the Department of Molecular Biology at the Research Institute of Scripps Clinic. He spent nine years in the Chemistry Department at the University of California, Davis, before moving to his present position in 1986. His research interests include dynamical simulations of proteins and nucleic acids and electronic calculations of active sites of metalloenzymes.

References

1. Wuthrich, K., *Science*, Vol. 43, pp. 45-50, 1989; Wright, P. E., *Trends in Biochemical Science*, Vol. 14, pp. 255-260, 1989.
2. Havel, T. F., I. D. Kuntz, and G. M. Crippen, *Bulletin of Mathematical Biology*, Vol. 45, pp. 665-720, 1983.
3. Brünger, A., "Crystallographic Refinement by Simulated Annealing on Supercomputers," *CRAY CHANNELS*, Vol. 10, Number 3, pp. 16-19, Fall 1988.
4. Yip, P. and D. A. Case, *Journal of Magnetic Resonance* 83, pp. 643-648, 1989.
5. Lee, M. S., G. P. Gippert, K. V. Soman, D. A. Case, and P. E. Wright, *Science*, Vol. 245, pp. 635-637, 1989.
6. Moore, J. M., G. P. Gippert, D. A. Case, and P. E. Wright, unpublished results. See also Moore, J. M., D. A. Case, W. J. Chazin, G. P. Gippert, T. F. Havel, R. Powls and P. E. Wright, *Science*, Vol. 240, pp. 314-317, 1988.



Vorticity contours on the inside surface of the artificial heart chamber.

Numerical simulation of flow through an artificial heart

Stuart Rogers, Paul Kutler, Dochan Kwak, and Cetin Kiris
NASA Ames Research Center, Moffett Field, California

Advances in computational fluid dynamics (CFD) and in supercomputer hardware enable researchers to simulate more complicated flow problems than previously had been possible, and this capability extends beyond the field of aerodynamics. For example, researchers at NASA Ames Research Center are applying a newly developed solution algorithm to the problem of flow through a model of an artificial heart. Developing this capability will extend the use of CFD as a design tool to all varieties of fluid dynamic mechanisms. The calculations described here were performed on the CRAY-2 computer system at the NASA Ames Research Center.

Flow problems

Medical researchers designing prosthetic devices for use in the circulatory system need to understand the fluid dynamics of blood flow through the heart and blood vessels. An understanding of blood flow is necessary to design prosthetic devices, such as artificial hearts, because turbulent flows create shear stresses that can damage blood cells, and stagnation in a flow promotes clotting. A thorough understanding of blood flow requires detailed knowledge of flow quantities. Blood may exhibit significant non-Newtonian characteristics locally and the geometries involved in the circulatory system usually are very complicated. The flow also is unsteady, possibly periodic, very viscous, and incompressible. The problems involved in modeling this type of flow are highly interdisciplinary, and an attempt at a complete solution would be a formidable task. However, an analysis based on a simplified model may provide much needed physical insight into blood flow analysis. References 1 to 4 provide a more comprehensive background on blood flow research.

Mechanical hearts increasingly are in demand as temporary life support systems. Presently, these devices have several problems, many of which are attributable directly to the fluid dynamics of the blood flow within them. Medical researchers therefore stand to benefit greatly from the application of state-of-the-art computational fluid dynamics technology to determine flow characteristics in these devices. Ongoing work at the NASA Ames Research Center has involved the development of viscous, incompressible flow solvers. This research is motivated by the need for realistic three-dimensional simulations for aerospace applications, such as the flow through the Space Shuttle main engine power head.

The formulation described in this article is based on a Newtonian fluid assumption. However, because the governing equations are solved in a generalized coordinate system, viscosity that varies in space and time is allowed; a full simulation of viscoelastic flow is very difficult to perform because of the fluid's nonlinearities. However, as a first step toward full simulations, non-Newtonian effects of the blood flow can be simplified by a constitutive model for viscous stresses. In addition, the code formulation allows the implementation of a moving geometry. Thus, these flow solvers can be applied to the analysis of mechanical hearts and ventricular assist devices. The primary purpose of this work is to transfer NASA-developed technology to artificial heart researchers. NASA benefits by advancing the state of the art in CFD for treating unsteady internal flow with moving boundaries. The artificial heart manufacturers benefit by gaining a better understanding of the fluid flow within their devices and, ultimately, better device designs.

Computational solution method

Recent developments in the numerical solution of the incompressible Navier-Stokes equations include an efficient algorithm for time-dependent flows.[5] This algorithm uses a flux-difference-split, upwind-differencing scheme for the convective fluxes. The resulting system of numerical equations is more nearly diagonally dominant than that resulting from the use of a central-difference scheme. The system is solved using an unfactored line-relaxation scheme that has very good stability and convergence characteristics.

Geometry and grid generation

The actual model of the Pennsylvania State artificial heart, designed at Pennsylvania State University, poses some difficult problems from a computational standpoint. Figure 1 shows a computer-generated, shaded-surface representation of the heart. The heart

comprises a cylindrical chamber with two openings on the side for valves. The pumping action is provided by a piston surface that moves up and down inside the chamber. The diameter of the piston is 7.4 cm, with a stroke length of 2.54 cm. The problem was nondimensionalized with a characteristic length of 2.54 cm and a characteristic velocity of 40 cm/sec. The actual artificial heart has a cylindrical tube extending from each of the valve openings. These tubes contain tilting flat disks that act as valves. The computational model described here neglects the valves altogether and uses the right and left openings shown in Figure 1 for the inflow and outflow boundaries, respectively. In the computations, as the piston reaches its highest position, the outflow valve closes and the inflow valve opens instantaneously. Similarly, as the piston reaches its lowest position, the outflow valve opens and the inflow valve closes.

In the actual heart device, the piston moves through the entire chamber volume, including across most of the valve opening. This causes some severe problems for the computational grid as the piston moves across the valve boundaries. Because the flow solver is designed to use body-fitted coordinates, grid lines must be placed around the valve to coincide with the valve opening boundaries. But because the piston moves past this opening, the grid must accommodate both of these surfaces. A computational grid can handle this motion in several ways. One method involves a Chimera scheme in which two grids are used, one that moves with the piston, and one that is attached to the rest of the body.[6] Information is passed between the two grids by interpolating variables at the grid interfaces. This method can be expensive and complicated to implement.

Two methods that are somewhat simpler to implement both involve the use of one grid inside the computational domain. One of these involves a stationary grid through which the piston surface travels. Boundary conditions applied at the piston surface would ensure that mass and momentum are conserved for the partial grid cells as the piston moves through them. However, with the curved surfaces at the valve boundaries, this would involve partial grid cells with an arbitrary number of sides, making this formulation difficult. Another single-grid method uses a grid that varies in time as the piston moves through the boundary. The generalized coordinate transformation used in this method enables the grid motion to be accounted for in the equations with the time-varying metrics.

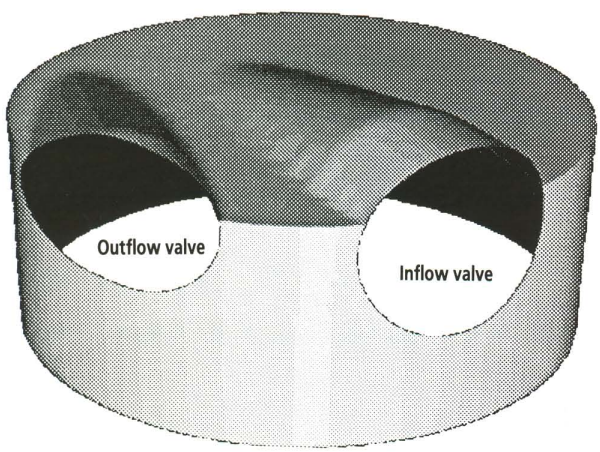


Figure 1. Artificial heart geometry showing valve openings.

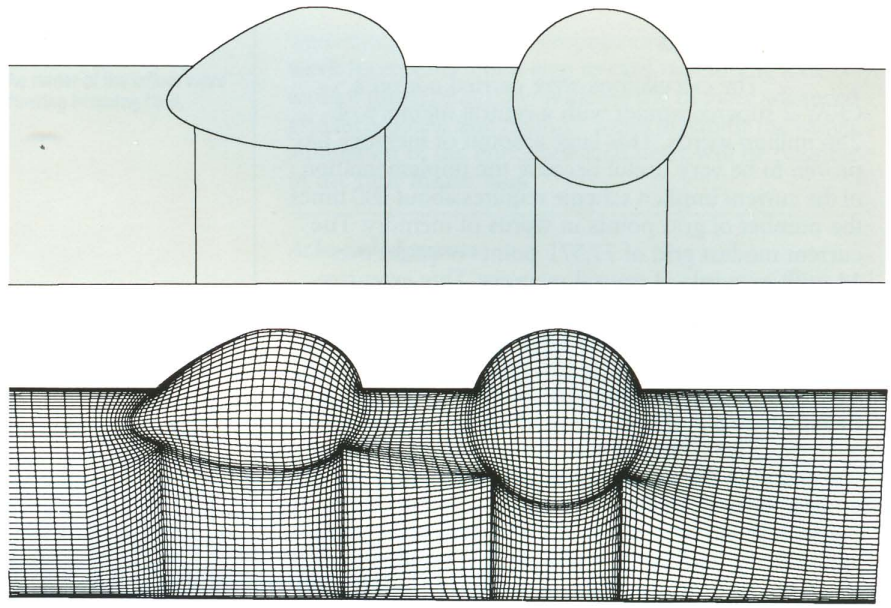


Figure 2. Unwrapped side surface of the artificial heart: (top) zone boundaries showing the surface sections used for grid generation, (bottom) completed surface grid.

This scheme requires that a separate grid be generated for each discrete position of the piston during the calculations. This may provide the simplest way to accommodate the entire piston motion, although it poses a difficult grid-generation problem.

To eliminate this problem, a simpler piston motion was chosen for the calculations of the present study. The piston was restricted from rising above the bottom of the valve openings. The latter of the single-grid approaches was used, so that a constant number of grid points were used and the grid between the piston and the bottom of the valve openings compressed and expanded as the piston moved up and down. This was achieved by allowing the piston to travel farther down so that the overall volume of the chamber in the computations was larger than in the actual device. One possible drawback to this approach is that the continuous compression and expansion of the grid will have a varying effect on the accuracy of the flow simulation because the grid density will change with time. Any such effect has yet to be studied in detail.

To make the most efficient use of grid points, an H-H grid topology was used to fit the grid to the physical domain. The grid dimensions were chosen to be 39 by 39 by 51 because of memory and computational time limitations of the current flow solver. To generate a grid at each time step for this geometry, the surface grid was generated first, and from that an algebraic grid generator and elliptic smoother were used to generate the interior points using the distribution given on the surface grid. To generate the surface grid, the side boundary was divided into seven zones. The points were distributed along each of the zonal boundaries, and then a biharmonic solver was used to generate the grid interior to each of the zones. The biharmonic solver also was used to generate an H-grid for the top and bottom surfaces of the heart device. This approach made it relatively simple to repeat the process at each time step for any given position of the piston surface. Figure 2 shows the unwrapped surface of the side of the heart chamber. The top image shows the zonal boundaries of the different sections used to generate the surface grid, and the bottom image shows the resulting surface grid.

Computed results

The calculations were carried out on a CRAY-2 supercomputer with a central memory of 256 million words. This large amount of memory has proven to be very useful because the implementation of the current implicit scheme requires about 180 times the number of grid points in words of memory. The current modest grid of 77,571 points requires over 14 million words of central memory. This extensive memory use was necessary because the line-relaxation scheme was coded in such a way as to save computational time by storing terms and utilizing more memory whenever possible. Other formulations using less memory are possible and will be coded in the future.

The computations started with the fluid at rest and with the piston at the bottom position and the outflow valve open. The computations were carried out for a Reynolds number of 100 based on unit length and velocity, and the flow was assumed to be laminar. In the actual heart the Reynolds number is about 600, and regions of the flow are turbulent. The laminar assumption is used here because the main purpose of this calculation is to test the ability of the flow solver to compute flow through this complicated geometry, separate from the effects of an inadequate turbulence model. Finally, the fluid is assumed to be Newtonian, which corresponds to the experiment of Tarbell[3], who used a water and glycerin fluid with a viscosity similar to that of blood, about 3.5 centipoise. Unlike blood, however, the fluid exhibits a Newtonian fluid behavior.

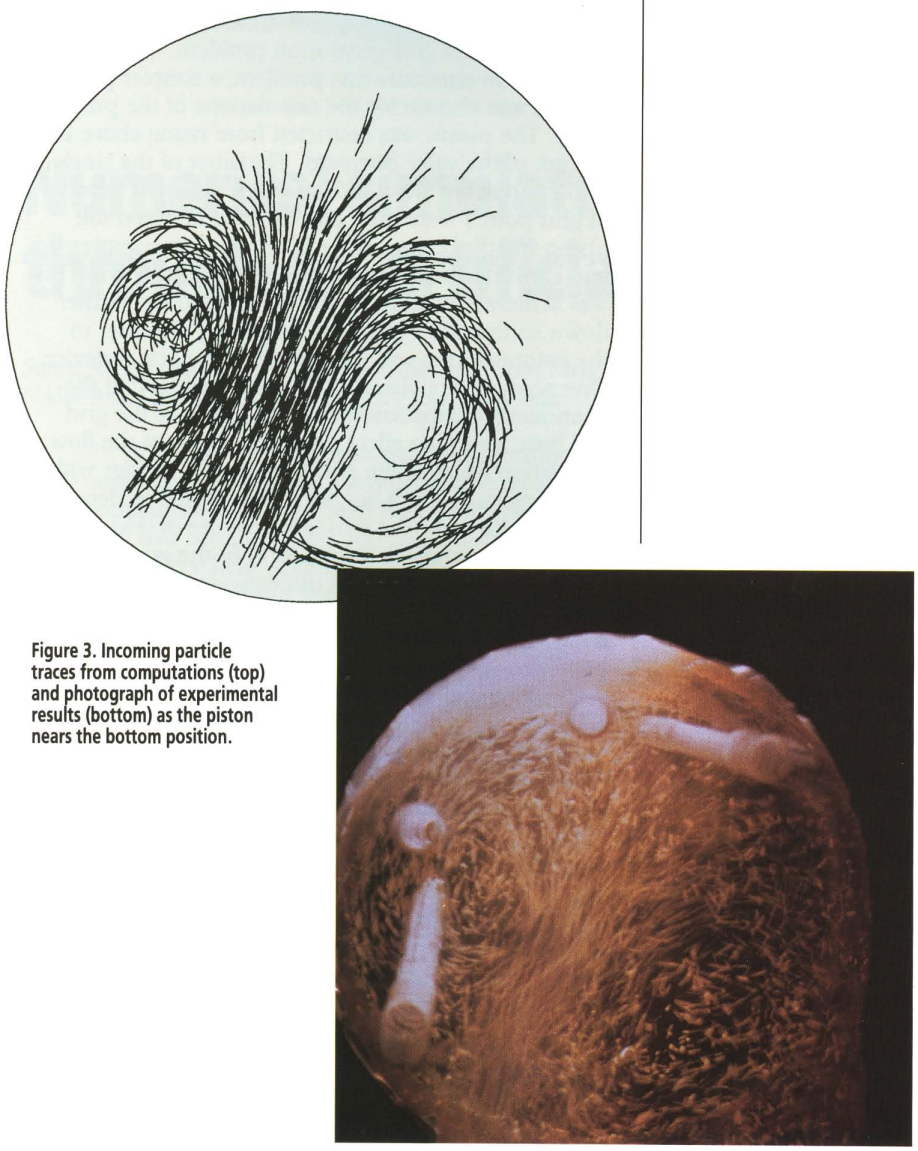


Figure 3. Incoming particle traces from computations (top) and photograph of experimental results (bottom) as the piston nears the bottom position.

The use of characteristic relations[5] at the inflow and outflow boundaries determines only part of these boundary conditions. At the inflow, three variables must be held constant, and at the outflow one variable must be held constant. At the inflow valve opening, the total pressure is specified to be constant and the velocity is prescribed to be perpendicular to the open surface. At the outflow valve opening, the static pressure is specified to remain constant. This scheme provides a nonreflective boundary treatment that remains computationally stable. The rest of the boundaries are prescribed to be viscous, no-slip surfaces, at which the pressure on the boundary is computed by specifying that the normal pressure gradient be zero.

The artificial compressibility constant β was set to 500, a value obtained from numerical tests to determine the best convergence during the subiterations. The larger the pseudotime-step $\Delta\tau$, the faster the convergence, and therefore $\Delta\tau$ was set to 10^{12}, which effectively set the $1/\Delta\tau$ term to machine zero. The physical time step Δt was set to 0.025 (values of Δt larger than this tended to make the computation slightly unstable). The piston moved with a constant, nondimensionalized velocity of ± 0.2 between its top and bottom positions, requiring 200 physical time steps for one period of the piston's motion. During each time step, the subiterations were carried out until the maximum residual dropped below 10^{-3} or until a maximum of 20 subiterations was used. During most of the piston's cycle only 12-15 subiterations were required, but when the piston was changing directions, it did not completely converge in 20 subiterations. This did not cause any stability problems, yet it remains to be seen what effect this has on the accuracy of the solution. The computing time required for each period of the piston's motion was approximately four hours. The computations were run for four periods during which particle paths were computed after being released near the inflow valve.

The top image in Figure 3 shows some of these particle traces as the piston nears its bottom position. Two distinct vortices are seen to have formed from the flow separating as it enters through the inflow valve and encounters the lower pressure regions adjacent to the valve. The bottom image in Figure 3, an experimental photograph (J. M. Tarbell, private communication, 1988), shows bubbles entering the inflow valve as the piston nears its bottom position. A similar two-vortex system is seen to form here. Figures 4 and 5 show velocity vectors during the inflow phase in planes passing through the center of the inflow valve. Figure 4 shows a top view of vectors in a plane parallel to the piston, while Figure 5 shows a side view of vectors in a plane perpendicular to the piston. These figures portray the complexity of the vortical structure of this flow. Figure 4 again shows the presence of two vortices formed as the incoming flow forms a jet. Figure 5 also shows how the flow recirculates underneath the valve opening, but it does not show that

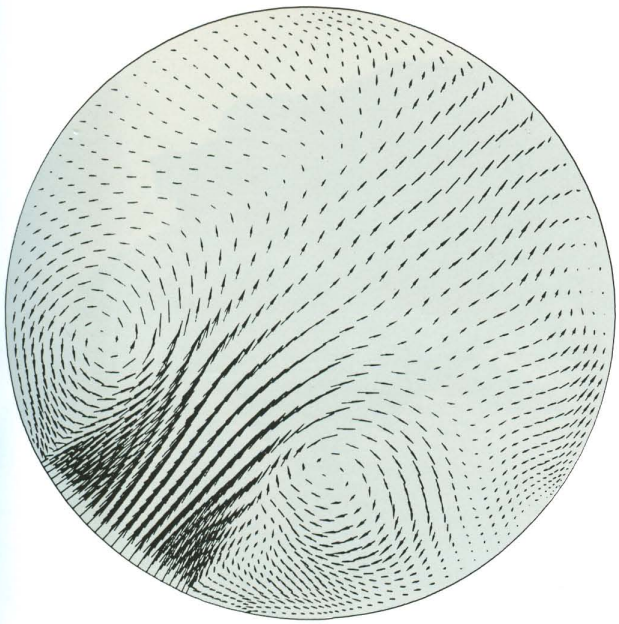


Figure 4. Top view of velocity vectors in a plane through the center of the inflow valve showing incoming fluid.

the flow there is strongly three-dimensional, with the velocity vectors next to the left wall underneath the valve pointing into the paper. The figure also shows the presence of additional vortices or stagnant flow regions against the back wall opposite the valve opening. This is a region that possibly could benefit from a design modification.

Conclusion

An algorithm for computing unsteady incompressible Navier-Stokes equations has been extended to simulate flow through an artificial heart. The present solution shows the capability of the computational procedure for simulating complicated internal flows with moving boundaries. For this initial calculation, the motion of the piston of the actual device was simplified. The fluid was assumed to be laminar and Newtonian. Also neglected were the effects of the valve opening and closing. Even though the computer code, which is based on a nonfactored implicit line-relaxation scheme, converged rapidly, further enhancement in computational efficiency still will be useful. One simple modification would be the use of a multigrid convergence-acceleration scheme. Many different configurations will need to be analyzed in the design of the heart, which will require a faster flow-solver code. One difficulty in performing the present study is that very little validation can be done because of the laminar flow assumptions and because the valves were simplified in the computation. Nonetheless, this work represents a first step toward developing a CFD tool for this type of flow. With the use of more advanced grid techniques, and multiple zones to handle the valves, simulation of the entire artificial heart will be possible in the near future. ▬

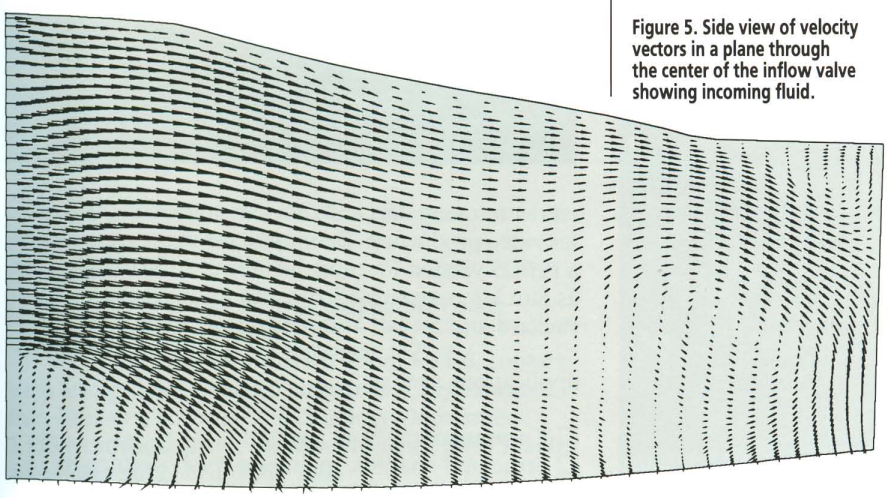


Figure 5. Side view of velocity vectors in a plane through the center of the inflow valve showing incoming fluid.

Acknowledgments

This research described in this article is part of a joint effort with Pennsylvania State University and Stanford University and is partially funded by the NASA Technology Utilization Office. This article appeared in an expanded form as NASA Technical Memorandum 102183.

About the authors

Stuart E. Rogers received B.S. and M.S. degrees from the University of Colorado, in 1983 and 1985, respectively. He received a Ph.D. degree in aeronautics and astronautics from Stanford University in 1989. He has been working at the NASA Ames Research Center for the past five years in the Applied Computational Fluids Branch, specializing in incompressible flow problems.

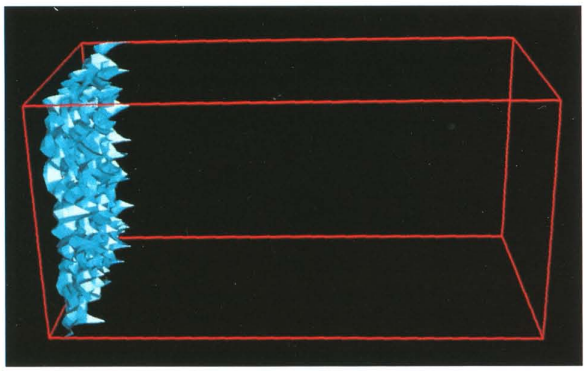
Paul Kutler received B.S., M.S., and Ph.D. degrees from Iowa State University in 1965, 1967, and 1969, respecively. He since has been working at the NASA Ames Research Center, primarily in computational fluid dynamics, and currently is chief of the Fluid Dynamics Division.

Dochan Kwak received a B.S. degree in 1964 from the Seoul National University in Korea. He received a Ph.D. degree in mechanical engineering from Stanford University in 1975. Until 1979, he worked at Los Alamos National Laboratory. Since then he has been working at the NASA Ames Research Center in the Applied Computational Fluids Branch, where he is leader of the incompressible computational fluid dynamics group.

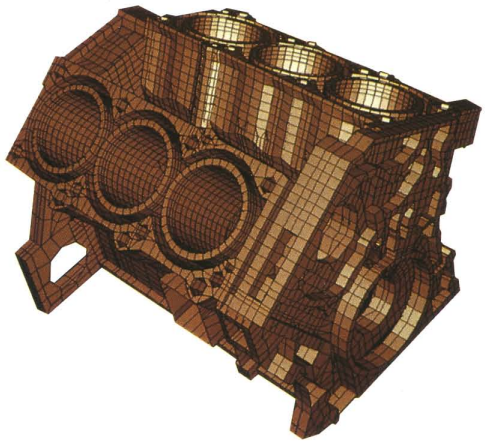
Cetin Kiris graduated from Istanbul Technical University in 1983. He received an M.S. degree from Stanford University in 1987. He currently is working on his Ph.D. degree in aeronautics and astronautics at Stanford University while also working at the NASA Ames Research Center through a consortium agreement.

References

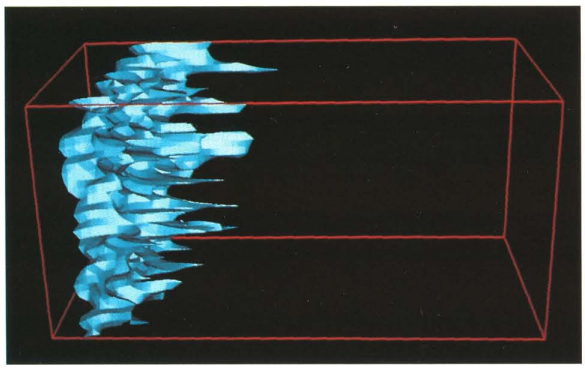
1. Tiderman, W. G., M. J. Steime, and W. M. Phillips, "Two-Component Laser Velocimeter Measurements Downstream of Heart Valve Prostheses in Pulsitile Flow," *Journal of Biomechanical Engineering,* Transactions of the ASME, Vol. 108, pp. 59-64, 1986.
2. Liepsch, D., H. J. Steiger, A. Pon, and H.-J. Reulen, "Hemodynamics Stress in Lateral Sacular Aneurysms," *Journal of Biorheology,* Vol. 24, pp. 689-710, 1987.
3. Tarbell, J. M., J. P. Gunshinan, D. B. Geselowitz, G. Rosenburg, K. K. Shung, and W. S. Pierce, "Pulse Ultrasonic Doppler Velocity Measurements Inside a Left Ventricular Assist Device," *Journal of Biomechanical Engineering,* Transactions of the ASME, Vol. 108, pp. 232-238, 1986.
4. Peskin, C. S., "The Fluid Dynamics of Heart Valves: Experimental, Theoretical and Computational Methods," *Annual Review of Fluid Mechanics,* Vol. 14, pp. 235-259, 1982.
5. Rogers, S. E., and D. Kwak, "Numerical Solution of the Incompressible Navier-Stokes Equations for Steady-State and Time-Dependent Problems," AIAA Paper 89-0463, 1989.
6. Beneck, J. A., P. G. Buning, and J. L. Steger, "A 3-D Chimera Grid Embedding Technique," AIAA Paper 85-1523-CP, 1985.



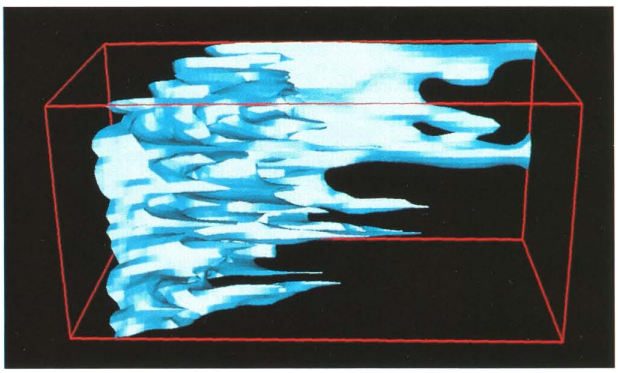
Left and below, image sequence from a viscous fingering model from British Petroleum. The images show the location of the 50 percent oil/water barrier in an oil field at various times following water injection.



Right, MSC/NASTRAN model of a V6 engine block from General Motors. The model was used for vibration analysis.



Supercomputer graphics for engineering applications



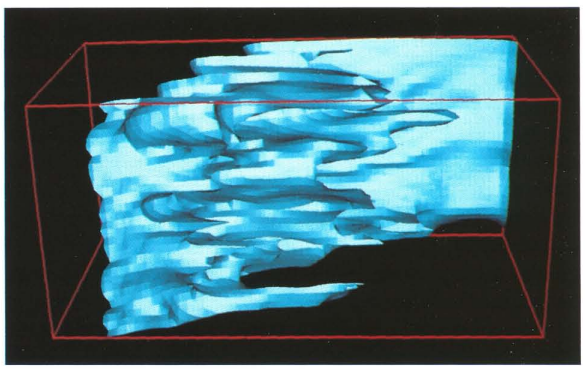
MPGS

The Cray Research Multipurpose Graphics System (MPGS) is an interactive menu-driven visualization tool for use on Cray computer systems. It is the only advanced-graphics visualization package written to support supercomputer applications. Using TCP/IP network protocol, MPGS is distributed between a UNIX-based workstation and any Cray system running Cray Research's UNICOS operating system.

The MPGS package offers true distributed processing, whereby the Cray system processes the computationally intensive data, and the workstation performs local graphics manipulations. This workload distribution ensures the efficient use of both computer systems and minimizes network data transfers. The required TCP/IP protocol also ensures that information can be transmitted between a Cray system and a user workstation over any distance.

Because MPGS brings to the workstation only small amounts of data at a time, powerful workstations and powerful communication channels are unnecessary. Using MPGS, data service centers can handle true graphics applications.

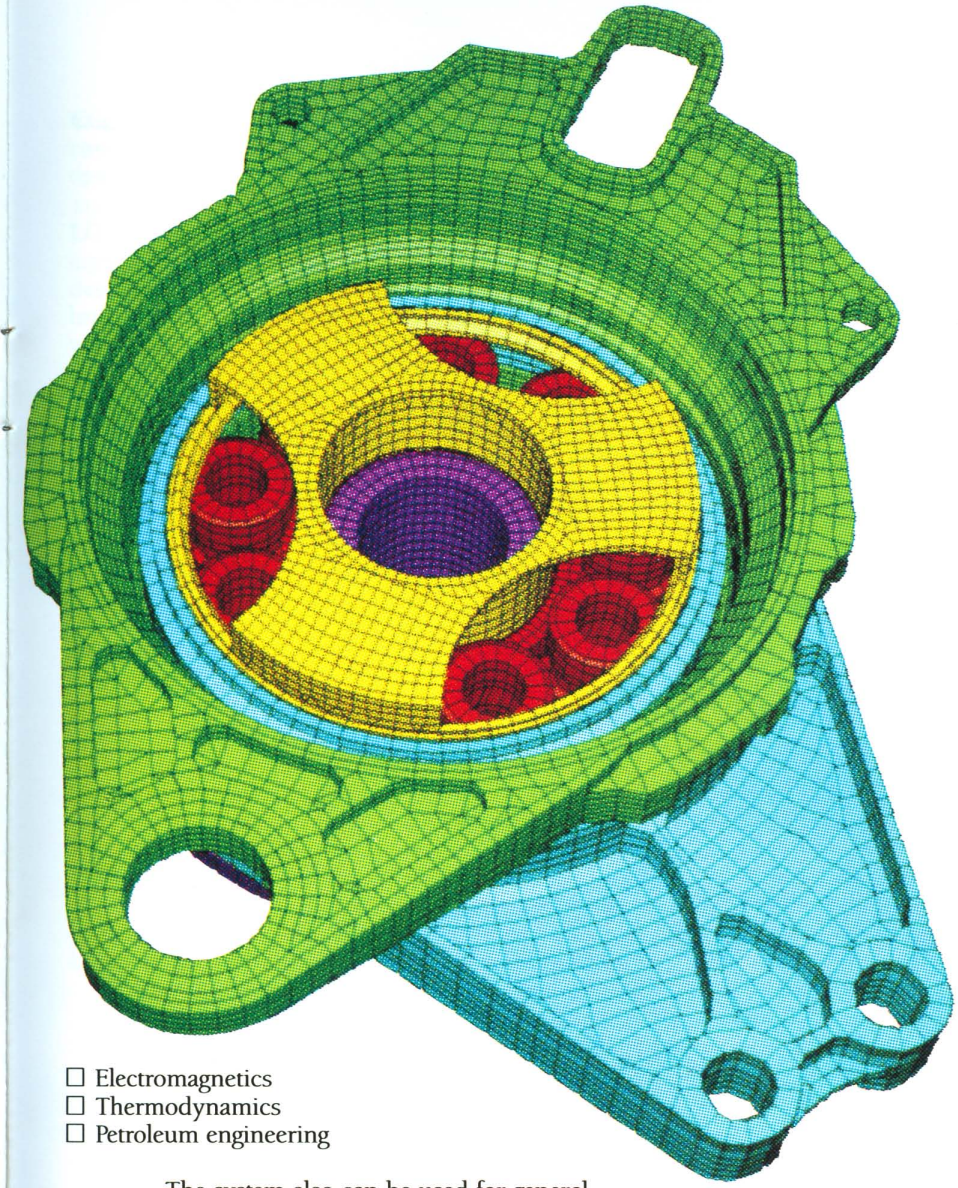
The MPGS graphics package runs on Silicon Graphics IRIS workstations, and the latest release can

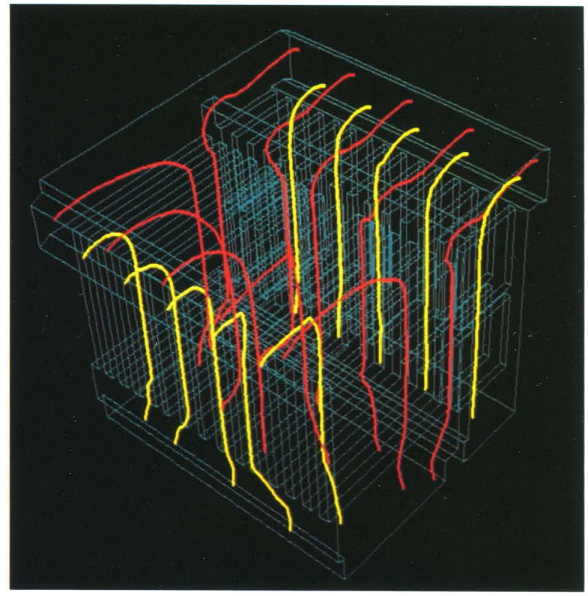


render transient data and provides animation capabilities for video production. Developments in progress include a PHIGS version for most high-end workstations; such as Sun Microsystems' Sparc and those from Hewlett-Packard and Stellar; and mesh generation.

MPGS is intended for use in engineering applications such as

- ☐ Computational fluid dynamics
- ☐ Structural analysis, including crash simulation





Above, FIDAP simulation of airflow through a Hewlett-Packard computer cabinet. The model was used for cooling studies.

Left, 33,619-element MSC/NASTRAN model of a rotary actuator from the Zahradfabrik Friedrichshafen Aerospace Division in West Germany. The actuator is for the flap activating system on an Airbus A330/A340 airplane. The image shows the five superelements used for the analysis of a point load.

Below, DYNA-3D model of an aircraft fan blade from Rolls Royce. The model contained approximately 250,000 hexahedron elements and demonstrates the ability of the MPGS package to handle very large problems.

- ☐ Electromagnetics
- ☐ Thermodynamics
- ☐ Petroleum engineering

The system also can be used for general graphics rendering and meteorology applications.

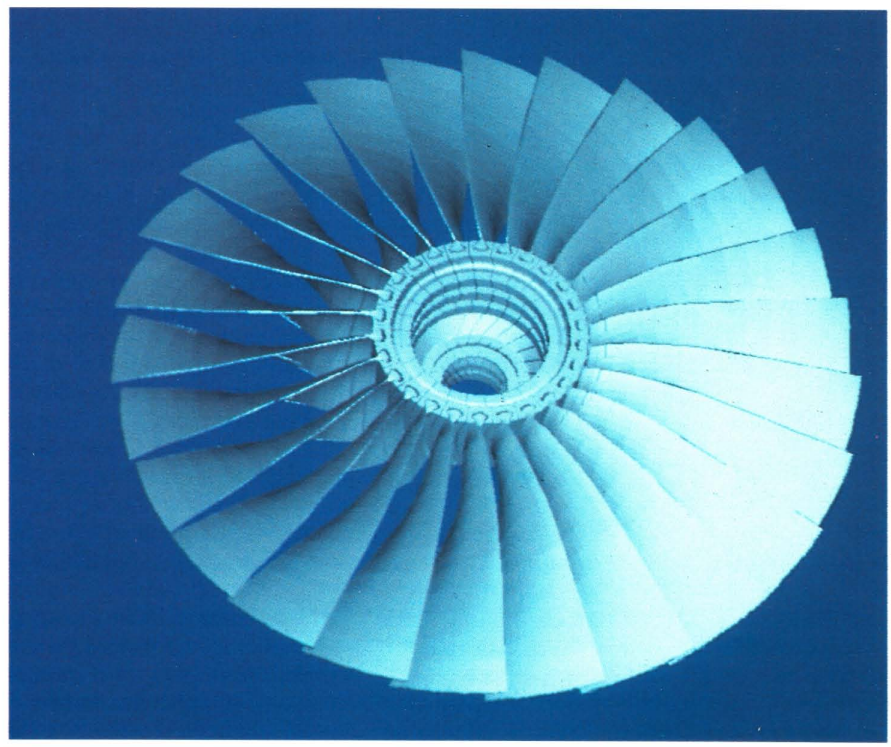
MPGS also handles many types of data structures, including

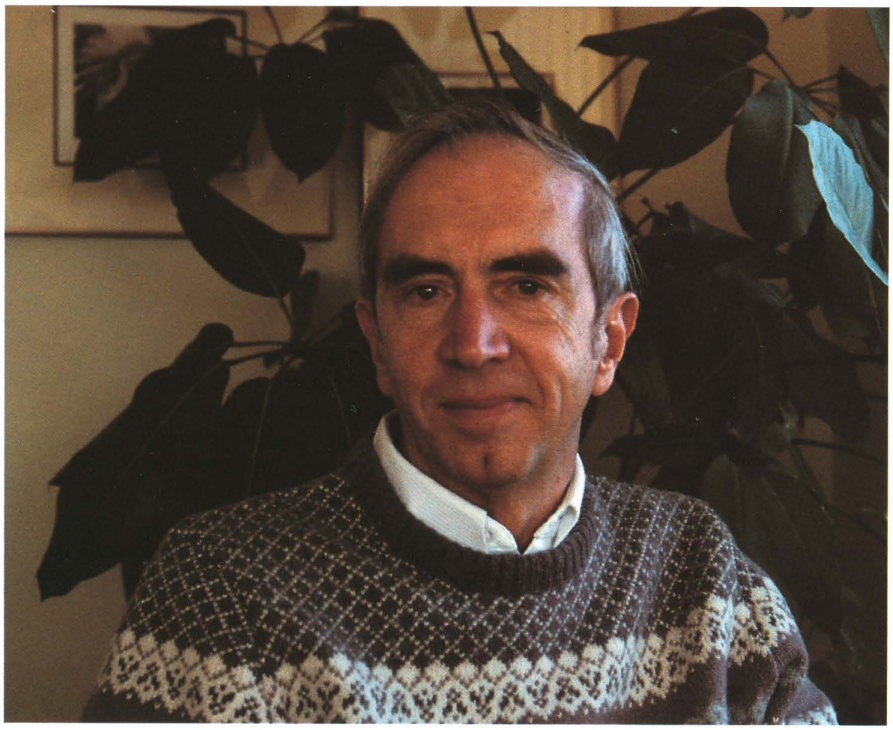
- ☐ Finite element
- ☐ Finite difference
- ☐ Finite volume
- ☐ Boundary element

MPGS enables the visualization of results from many engineering application programs, including

- ☐ MSC/NASTRAN
- ☐ ABAQUS
- ☐ PAM-CRASH
- ☐ DYNA3D
- ☐ RADIOSS
- ☐ PATRAN
- ☐ FIDAP
- ☐ FLUENT/BFC and most aerodynamic codes

For more information about MPGS, contact Kent Misegades or Anders Grimsrud, Cray Research, Inc., Industry, Science & Technology Department, 1333 Northland Drive, Mendota Heights, MN, 55120; telephone: (612) 681-3660.





Commitment to performance

An interview with Les Davis

Lester T. Davis is executive vice president of Chippewa Falls operations for Cray Research. Davis, a founder of Cray Research, was the chief engineer for the CRAY-1 project. He became a vice president of the company in 1973. Prior to joining Cray Research, Davis was director of electrical engineering and general manager of the Chippewa Laboratory for Control Data Corporation. In this interview, he shares his thoughts on Cray Research's directions in hardware development, the company's technical strengths, and general trends that will characterize supercomputing in the 1990s.

CC: The CRAY Y-MP computer system is Cray Research's mainstay product today. The original system was introduced in 1988, and the full series has been on the market for about a year. What kind of feedback have you received so far from CRAY Y-MP system users?

LD: The response from users and from customers has been overwhelmingly positive. In almost all cases they report that the systems have exceeded their expectations in terms of performance and reliability. We expect that these systems will continue to be well received, particularly as we add enhancements, such as larger memories and larger SSD storage devices. In fact, our enhancement plans for this product line probably are more aggressive than for any previous product line.

CC: The creation of Cray Computer Corporation to complete the CRAY-3 project, along with a growing market for low-end supercomputers, might lead some people to question whether Cray Research will, or can, retain its commitment to build and support the most powerful computers in the world. How would you respond to such a concern?

LD: You just have to look at the people who are working here at Cray Research and the kinds of projects that are underway to know that we are committed to maintaining our leadership position at the high-end. We have a team working on a follow-on to the CRAY Y-MP system, which will be a 16-processor system with two vector pipes per processor. The system should reach a peak performance in excess of 15 gigaflops. We already have a working floating-point multiply unit, a floating-point add unit, and several other significant test vehicles; so this project is well under way.

The CRAY Y-MP follow-on system will be an all-silicon machine, but we are exploring the use of gallium arsenide components in future systems. That potential and other ideas in component and packaging technologies are being looked at by our technology development group. We also have a research team looking at options for future architectures, such as massively parallel systems. At the same time, we're studying the environments that are needed to support high-end systems; here we're looking at networking technologies and at smaller systems that we can create by extending the CRAY Y-MP architecture at the low end. Low-end systems can do more to support the high-end in supercomputer environments and we want to provide that service to our customers.

Another form that our commitment to the high-end takes is the investment we've made in integrated circuit and printed circuit board manufacturing facilities. Having these and other facilities on site enables us to keep product development cycles as short as possible. Over the years we've also created a business and technical infrastructure that includes manufacturers of integrated circuits, networking equipment, disks, interconnects, and software; in some cases we have explicit licensing agreements with these people; in others, more informal relationships. This infrastructure also includes some of our large customers who help us come up with new product designs. Although Cray Research is interested in the final product — the high-performance systems — all of the people in this infrastructure want to participate in making these systems possible. This infrastructure gives us a strong position that helps us compete creatively and effectively with large vertically integrated companies. If we weren't committed to maintaining our lead in the high end, we wouldn't maintain all of these strategic relationships.

CC: Now that Seymour Cray no longer is a contractor to Cray Research, is the company short of design talent?

LD: Not at all. Seymour Cray was an outstanding individual among many who contributed to the design of Cray Research computer systems. But we have many young computer architects and designers working on various projects, and these are many of the same people who were central to the design and development of the systems that have brought the company its continued success. In recent years, we've added a number of new very talented people in all of the areas that are necessary to build these high-performance machines. We have a lot of strength in circuit design, architectural design, packaging, and manufacturing. And we have equally capable design teams working in software.

CC: To what extent do requests from customers influence the directions that Cray Research takes in product design?

LD: We consider customer input very valuable. We pay a lot of attention to our customers' requests. Our future architecture study team, in particular, has been talking to customers to assess what computing capacities industries will need in the future. If you look at it, all of our products have been built according to customers' requests; we always do that. Of course, we can't put everybody's wants into each product, but we can't design these systems in a vacuum either.

Some of our large customers have expressed an interest in massively parallel architectures, so we are looking at how massively parallel systems will play a role in the supercomputing environment. We don't imagine that our present strategy with parallelism — offering the fastest processors that the technology will allow and doubling the number with each major product — and massive parallelism are mutually exclusive. We're running some experiments now, in fact, to assess just what makes up a supercomputing environment. Is it just large numbers of general-purpose processors tied together? Probably not. Maybe it's that plus some kinds of special-purpose processors. And the special-purpose environment might be massively parallel processors plus some others. Generally, the kind of environment that we want to create is a general-purpose environment, although it might have more special-purpose character than it had in the past. And we'll need that to get the highest level of performance that we can during the 1990s.

CC: What are the most significant ways in which Cray Research systems differ from other supercomputers on the market?

LD: Well, the obvious point is performance. Our systems offer the highest throughput of any available computer systems. But there are some other important factors as well. We've come up with a strategy that allows customers to continue to build into the future; we don't offer our customers dead-end products. We take advantage of the latest technologies to build very reliable systems and yet deliver the highest performance that is possible with the technology. Today, our customers get the highest performance, they know the products have a future, and they have a commitment from Cray Research to proceed as aggressively as possible in the development of new hardware and software products.

CC: What interest does Cray Research have in exploring the low end of the supercomputer spectrum?

LD: We're committed to making the CRAY Y-MP architecture very competitive at the high and low ends. As far as the low end, we plan to extend the architecture downward as far as we can and remain competitive. That then, for us, will define the lower boundary of supercomputers. We don't see this as a distraction of resources from the high end, but rather as a way to complete the supercomputing environment, in which you might need a lower performance machine along with a high-end supercomputer. Technology is allowing us now, within the CRAY Y-MP family, to make the systems easier to install and maintain, with fewer recurring costs. Basically, we want to bring supercomputer power to the user at his or her desk; we want users to feel like they have a Cray system at their fingertips. The components necessary to provide this kind of environment will vary from site to site, and we want to retain the flexibility to meet a wide variety of hardware and software needs.

CC: What is the status of the CRAY-2 product line?

LD: Although we are not planning any enhancements for future CRAY-2 machines, the CRAY-2 series remains a fully supported product line. The CRAY-2 systems are very viable products, and we provide customers with an easy upgrade path from the CRAY-2 systems to the CRAY Y-MP line and an easy path will exist to move from the CRAY-2 to the CRAY Y-MP follow-on.

CC: What general trends in technology do you think will shape supercomputers during the 1990s?

LD: The machines of the future will not be determined by any one technology. You have to figure on some novel combinations of circuit technologies, packaging technologies, architectural concepts, and software design. And then you have to be able to build these machines in a cost-competitive environment. We can't ignore the fact that we're only interested in machines that can be built and that have a competitive price/performance — even though they have the highest performance in the world. Of course one way to help achieve this is to implement higher levels of automation in manufacturing.

I'm very optimistic about the 1990s. I see a lot of exciting opportunities for us in the marketplace, and we're in a good position to capture those opportunities. We now have a solid technical base working on a single product line. All of our resources are concentrated on creating the highest-performing systems and the proper environments to support them. I think the next decade will see an unprecedented growth in the demand for supercomputers, and we're positioned technically to remain the world's leading supplier. ▬

All of our resources are concentrated on creating the highest-performing systems and the proper environments to support them.

CORPORATE REGISTER

Cray Research lands new orders in the aerospace industry

Lockheed Aeronautical Systems Company (LASC), a subsidiary of the Lockheed Corporation, has installed a two-processor CRAY Y-MP computer system and peripheral equipment at the LASC Engineering Research and Development Center in Rye Canyon, California. The new system supplements a two-processor CRAY X-MP system, which was installed in the fourth quarter of 1985. The supercomputers are used for aeronautical systems engineering. LASC uses Cray Research's UNICOS operating system, which is based on AT&T UNIX System V. Doug Ford, Engineering Research and Development Center director, said, "Acquisition of the CRAY Y-MP system will allow us to meet our commitments, and to continue with the development and exploitation of advanced technology analyses and simulation methodologies."

Grumman Data Systems, prime contractor for NASA's Engineering Computation Facility at **Johnson Space Center** (JSC) in Houston, Texas, recently installed a CRAY X-MP EA/464 computer system. This is the first installation of a Cray system at JSC, which is NASA's mission control center and headquarters for astronaut training. JSC will use the Cray system for structural analysis, thermodynamics, and computational fluid dynamics to support NASA's space programs. For example, the supercomputer will be used to define expected loads, to determine flow fields around launch vehicles, and to calculate wind loads on launch days to provide timely and accurate launch decisions. The supercomputer will run Cray Research's UNICOS operating system.

Aérospatiale, a French aerospace company, has installed a CRAY X-MP/116se supercomputer at its new computer center in Les Mureaux, France. The computer is used by Aérospatiale's Space and Strategic Systems Division, a new customer for Cray Research, for fluid dynamics, structural analysis, and electromagnetics research and design. The company uses Cray Research's UNICOS operating system. Aérospatiale Space and Strategic Systems Division designs and manufactures satellites and space vehicles, such as the European space plane Hermes and the civil launcher Ariane. Aérospatiale installed a CRAY X-MP/14se computer system in 1987 at its facility in Toulouse, France.

The **Toyota Motor Corporation** has ordered an eight-processor CRAY Y-MP system, subject to export license approval. The purchased system is scheduled for installation in the first quarter of 1990 at Toyota's headquarters in Aichi, Japan. The company will use the system for engineering applications related to automotive design and development, including structural analysis, crash simulation, aerodynamic analysis, and design optimization. Toyota will use Cray Research's UNICOS operating system. This will be the second Cray system for Toyota, which installed a CRAY X-MP/18 system in October 1988.

Commissariat à l'Energie Atomique (CEA), the French Atomic Energy Commission, has installed a CRAY-2/4-256 computer system at its computer facility in Grenoble, France. The system, which is part of CEANET, the communication network of the French Atomic Energy Commission, is used for scientific research and development and supplements existing Cray systems installed at other CEA facilities.

The United Kingdom Meteorological Office, a new customer for Cray Research, plans to install an eight-processor CRAY Y-MP computer system with SSD solid-state storage device. The system is the first to be purchased through a joint marketing agreement between Cray Research and Control Data Corporation. The CRAY Y-MP supercomputer is scheduled for installation in the first quarter of 1990 at the headquarters of the Meteorological Office in Bracknell, Berkshire, U.K. The Meteorological Office will use the system for numerical weather prediction in support of day-to-day weather forecasting and for climate research. Currently, nine Cray Research supercomputers are installed around the world for weather forecasting and environmental research. Seventeen Cray Research systems are in use today in the United Kingdom.

Cray Research offers release 1.1 of the Cray Ada Environment

Release 1.1 of the Cray Ada Environment provides a set of tools and libraries that facilitate flexible, project-oriented software development. It also enables users to call routines written in Fortran, C, Pascal, and the Cray assembly language (CAL). Release 1.1 contains many enhancements to the two main components of the environment: the compiler and the debugger.

This release conforms to the ADA Compiler Validation Capability suite version 1.10 of the U.S. Department of Defense, thereby meeting an essential requirement of many United States government contracts. The Cray Ada Environment runs under the UNICOS operating system release 5.0 or later on CRAY Y-MP and CRAY X-MP EA systems in X-mode and on all CRAY-2 systems and on CRAY X-MP systems with Extended Memory Addressing. New compiler features include

- ☐ Support for the following Pragmas
  - — System_Name
  - — Storage_Unit
  - — Memory-Size
  - — Suppress_All
- ☐ More accurate handling of both fixed-point and floating-point numbers
- ☐ Conformity with the following standards
  - — Ada Issues Standard AI-00306 with regard to Pragma INTERFACE
  - — Ada Implementors Guide with respect to derived specifications
  - — Ada Implementors Guide and Ada Issue 00016 for expanding a name whose prefix is a renamed package
- ☐ Support for six implementation-defined attributes used for specifying type information
- ☐ Trackback information for UNICOS signals

New debugger features include

- ☐ More flexibility in setting and removing breakpoints; expanded breakpoint information
- ☐ Additional functionality to source- and machine-level instruction stepping
- ☐ An exception-trapping mechanism
- ☐ An option that allows for command-line invocation of debugger script files
- ☐ A user-defined debugger capability
- ☐ A history mechanism
- ☐ Multiple debugging windows
- ☐ Mnemonic task names
- ☐ User-defined macros and debugger variables
- ☐ Enhancements to source-level variable display and modification
- ☐ Support for Ada structured-programming constructs
- ☐ Type and representation display of objects

For more information about this release of the Cray Ada Environment, contact the nearest Cray Research sales office.

APPLICATIONS UPDATE

Petroleum reservoir simulation at one GFLOPS

In recent tests, the UTCHEM petroleum reservoir simulator ran faster than one GFLOPS (one billion floating-point operations per second) on a dedicated eight-processor CRAY Y-MP computer system, using Cray Research's Autotasking software feature. UTCHEM is a three-dimensional compositional chemical flood simulator developed at the University of Texas at Austin. Engineers in the Department of Petroleum Engineering optimized the code so that the latest version is about 95 percent vectorized. The code uses a variation of the IMPES method, and the solver is Jacobian Conjugate Gradient. The model used for this study was a reservoir divided into 30,976 (88 by 88 by 4) grid blocks, in the quarter five-spot geometry (one injector well and one producer well at diagonally opposite corners of a square horizontal area), containing residual oil filling about 35 percent of the pore space.

Without programmer-inserted directives, the code ran at 1.082 GFLOPS. Inserting directives into one subroutine to force parallelization of a large outer DO loop, which was not recognized by the Autotasking utility as safe to run in parallel, achieved an improved speed of 1.108 GFLOPS. Wall-clock timings indicate parallelization speedups of 5.51 and 5.65, respectively, over the original code on one CPU. The runs were made on a CRAY Y-MP system with a 6.41 nsec clock at Cray Research's computer center in Mendota Heights, Minnesota. Proportionately faster speeds are expected for systems with 6.00-nsec clocks.

For this particular enhanced oil recovery process, surfactant, polymer, and water are injected into a reservoir to free and produce the trapped oil. Simulation is extremely important in planning enhanced oil recovery processes because the operations are difficult and expensive to perform, and success depends highly on achieving an optimum combination of many parameters. Simulating such operations involves the solution of highly nonlinear partial differential equations that describe the flow of many fluid components. Simulations are complicated by complex chemical interactions between components. Both a fine grid and very small time steps are required for an accurate solution. Field-size problems are feasible only on supercomputers. The GFLOPS results reported here indicate that the CRAY Y-MP system can run these programs in reasonable time frames.

For more information on the UTCHEM simulator and its performance, contact Gene Shiles, Cray Research, Inc., Industry, Science & Technology Department, 1333 Northland Drive, Mendota Heights, MN, 55120; telephone: (612) 681-3636 or Gary Pope or Kamy Sepehrnoori, Department of Petroleum Engineering, University of Texas, Austin, TX, 78712; telephone: (512) 471-0231.

Convolution routine attains GFLOPS performance on CRAY Y-MP systems

Linear convolution is a computational technique commonly used in digital signal processing. Typical application areas include seismic exploration, geophysical research, crystallography, vibration analysis, radar and sonar signal processing, communications, speech recognition and synthesis, and the processing of weather data for analysis and prediction. A linear convolution routine recently developed at Cray Research has achieved a performance of two GFLOPS on an eight-processor CRAY Y-MP system while solving a one-dimensional convolution. This level of performance was achieved by using the Autotasking capability of Cray Research's CFT77 Fortran compiler. The routine was written completely in Fortran. The same code ran on a single-processor CRAY Y-MP system at 305 MFLOPS.

Given two sequences, the data sequence, $D = [d(i), i = 0, \ldots, n - 1]$, and the filter sequence, $G = [g(j), j = 0, \ldots, m - 1]$, where n is the data block size and m is the filter block size, the linear convolution creates a new sequence called the signal sequence, or the output sequence, $S = [s(k), k = 0, \ldots, m + n - 2]$. This sequence is given by the equation as $s(k) = \text{SUM}\,[g(k - i) * d(i)]$, where $i = 0, \ldots, n - 1$; $k = 0, \ldots, m + n - 2$; $g(k - i) = 0$ if $k - i$ is less than 0, and $m + n - 2$ is the output blocksize. From the equation, the number of floating-point operations for the linear convolution is 2 * (filter size) * (data size). Although an algorithm exists to compute linear convolution using a Fourier transform, the direct computation method still is used for many cases, as it is in the convolution routine described here. The table lists the performance of the routine for various filter lengths on an eight-processor CRAY Y-MP system. The routine also can be used to solve multidimensional convolution and correlation problems. For more information about this routine, contact Kuo Wu, Cray Research, Industry, Science & Technology Department, 1333 Northland Drive, Mendota Heights, MN, 55120; telephone: (612) 681-3682.

Data length	Filter length 5	10	20	50	100	200	500
500	107.62	122.33	131.64	138.26	161.31	176.52	170.48
1000	213.85	269.62	296.56	310.96	317.68	319.93	322.33
2000	337.25	437.32	493.90	521.11	534.84	540.73	545.89
5000	528.92	721.75	849.27	933.14	973.08	985.89	1000.88
10,000	662.07	915.58	1112.30	1251.03	1332.83	1362.21	1384.16
20,000	755.63	1070.32	1313.45	1548.73	1653.50	1705.81	1704.27
50,000	831.53	1186.54	1501.58	1783.66	1939.74	1958.77	2011.72

Performance in MFLOPS of the linear convolution routine on an eight-processor CRAY Y-MP system.

USER NEWS

Du Pont chemist honored as pioneer in theoretical chemistry

David A. Dixon, a Du Pont scientist who has pioneered the use of supercomputers in computational chemistry, has received the 1989 Leo Hendrik Baekeland Award. The biannual award, presented by the North Jersey Section of the American Chemical Society, is sponsored by Union Carbide and commemorates the inventor of Bakelite, the first synthetic plastic. Cray Research also was a corporate sponsor of the program.

Dixon was cited for leadership in theoretical chemistry; his work has opened new research areas in fluorocarbons, weakly bonded molecules, and main group element chemistry. The results from his work in computational chemistry increasingly are being used in place of physical experiments.

Two of the last four Baekeland award winners were computational chemists, notes Dixon, adding, "This shows that computational chemistry has become an important method with wide applications in solving chemical problems. The main tool for my research is Du Pont's Cray supercomputer. It has enabled me to solve real molecular problems in reasonable amounts of time."

After receiving a Ph.D. degree in physical chemistry at Harvard University in 1975, Dixon was appointed a junior fellow at Harvard. In 1977 he became assistant professor at the University of Minnesota and held that position until 1983 when he joined Du Pont's Central Research and Development Department.

Dixon plans to continue to push the field of computational chemistry, developing techniques to design molecules via numerical simulation. Currently, he is studying a wide range of polymeric systems and exploring substitutes for chlorofluorocarbons.

Researchers break performance barriers

Cray Research honored 20 researchers and research teams who are advancing the science of supercomputing and their fields of expertise at the 1989 Gigaflop Performance Award Ceremony held at the IEEE Supercomputing '89 conference in Reno, Nevada, on November 13. Award recipients solved real-world problems using Cray Research supercomputers running faster than one billion floating point operations per second (GFLOPS). Their achievements place them in a unique category of scientists and researchers at the forefront of progress.

Supercomputer performance not only enables scientists to solve in minutes problems that once took hours, but also enables them to look at problems in an entirely new light. For example, GFLOPS performance can enable researchers to examine data in three dimensions rather than in two. Effective use of parallel processing is enabling researchers to obtain results in reasonable times for models previously considered intractable.

While the applications had to run at sustained speeds of at least one billion floating point operations per second to be eligible for an award, the key was that the work be useful. Programs designed merely to put the supercomputer through its paces at high speeds were not accepted. Performance was defined as the total number of floating point operations for all central processing units divided by wall-clock time. This performance had to be reproducible. Nonstandard versions of compilers or other system software were not allowed.

The following summaries are just a sample of the outstanding work of Cray Research's Gigaflop Performance Award winners. To receive a brochure that describes the work of all the award winners, contact Diane Ciardelli, Marketing Communications, Cray Research, Inc., 1333 Northland Drive, Mendota Heights, MN, 55120; telephone: (612) 681-3630.

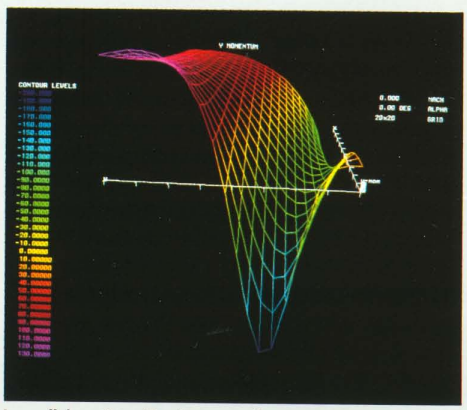
Demonstrating the potential for financial supercomputing

Determining the value of securities, making profits, and satisfying customers in the financial industry all depend upon making good decisions in limited amounts of time. Computers and sophisticated software enable traders to sift through reams of data and make complex calculations quickly, and to transfer processed information to customers on a timely basis. Unfortunately, many financial institutions are limited by conventional computer resources. Marjorie Hogan and Dexter Senft of The First Boston Corporation and Chris Moran of Cray Research recently ran a financial model on a Cray system to show that by using supercomputers, traders can have remarkably fast access to information about securities. When the model was tested on an eight-processor CRAY Y-MP supercomputer, the optimized code performed at a sustained rate of 1.39 GFLOPS.

The financial model was developed three years ago at First Boston to price options on bonds. A fixed income call option is the right to buy a bond at a fixed value, or strike price, during a specified exercise period. The fluctuation of interest rates and the passage of time, among other factors, determine the value of the option.

Certain fixed income securities have embedded options. Corporate bonds have the issuer's option to call the security from the investor. Mortgage-related securities have the homeowner's option to prepay principal. In each case, the value of the option affects the market value of the security. Traders of corporate bonds and mortgage-related securities need to value the embedded options in order to value the bonds.

The First Boston options model assumes that interest rates vary randomly and that



In a collaborative effort between The First Boston Corporation and Cray Research, analysts used a Cray supercomputer to study the characteristics of various fixed-income securities. The above graphic shows a three-dimensional view of the convexity characteristics of a callable corporate bond at various price and volatility levels.

options are exercised optimally. The computer model calculates the value of an option at thousands of interest rates and points in time, as it works backward from the expiration date to the present. In probability terms, the model calculates the probability-weighted value of the payoff, or the difference between the market value and the strike price on the exercise date. "The market trades securities with very complicated options," says Hogan. "The value of these securities is not intuitively obvious. Supercomputers can help traders price and trade unusual securities and help investors decide if and when they should buy them."

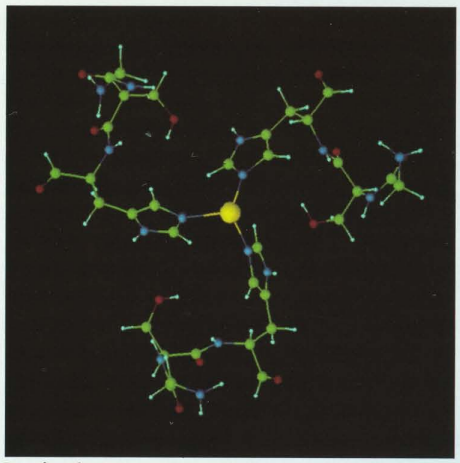
The demand for option valuation is enormous. "Whether we are figuring out how to market corporate bonds or calculating the value of embedded options in homeowners' mortgages, we have to price a lot of options all the time just to do our day-to-day business," says Hogan. "As it is, the traders may get one hundred inquiries a day and can run only 15 of them. This process could be much easier if traders and investors could run the model at will. A supercomputer could make this happen."

Study reveals details of insulin chemistry

In a collaborative project with researchers from public and private research institutions, Cray Research senior computational chemist Jan Andzelm achieved performance of 1.11 GFLOPS on a CRAY Y-MP system using the DGauss computational chemistry program. The program was written by Andzelm and uses the local spin density approximation and Gaussian-type orbitals to obtain wave functions and electron densities from Schrödinger's equation. The DGauss program implements direct sparse matrix algorithms that were designed at Cray Research for integral calculation, density synthesis, and numerical integration. The program achieves a high level of parallelism on Cray systems through the use of microtasking and Autotasking. As a result, electronic structures and charge distributions of large molecules can be calculated up to two orders of magnitude faster than with standard quantum chemistry methods.

Andzelm obtained GFLOPS performance with this code while investigating the binding of water molecules to the zinc ion in the model of human insulin. This research was conducted to elucidate the role of zinc in the crystallization of insulin. It is known that zinc plays a critical role in the formation of insulin crystals, but the process involved is not yet well understood.

The largest model considered during this work had nine amino acids (112 atoms), which amounts to 1031 basis functions at the 6-31G* level. Calculations for one geometry using this model required only 15 minutes of wall-clock time on a CRAY Y-MP system. This work revealed the precise positions of the water molecules, information that indicates uncertainty in corresponding experimental data. In addition, studies of the electrostatic potential surface around the zinc ion revealed other possible sites for weakly bonded water molecules.



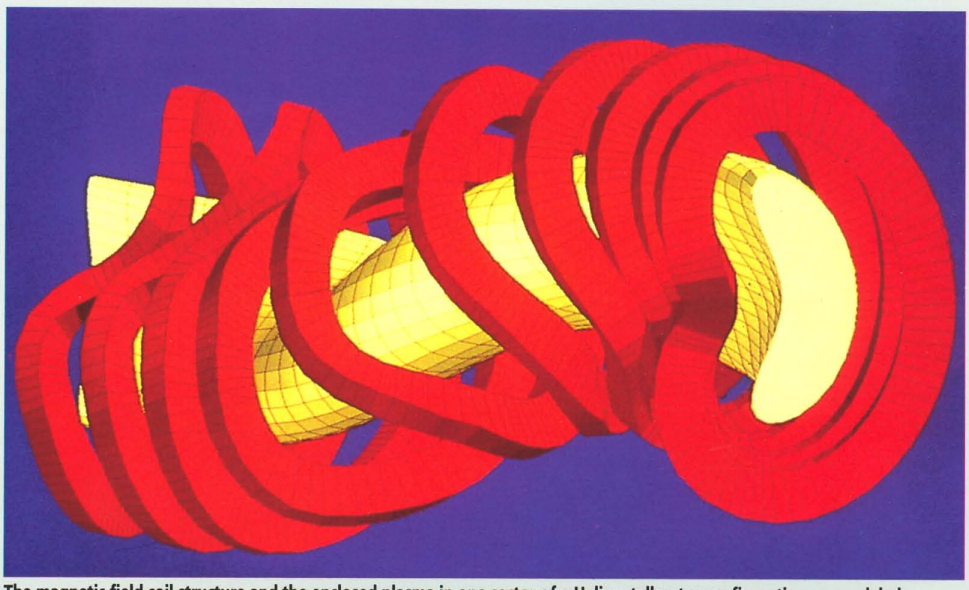
By using the DGauss computational chemistry program on Cray supercomputers, researchers can calculate electronic structures and charge distributions of large molecules up to two orders of magnitude faster than can be achieved with standard quantum chemistry methods. The above image shows a model of human insulin.

"Achieving this level of performance with this methodology enables us to calculate the charge distribution, electrostatic potential, binding energies, and geometries for molecular systems that are out of the reach of present quantum chemistry methods," Andzelm explains. "The methodology used in the DGauss program can achieve the functionality and accuracy of the standard ab initio quantum chemistry methods typically used in organic chemistry and biochemistry." Andzelm next plans to implement automatic gradient geometry optimization to further expand the functionality of the DGauss code on Cray Research computer systems.

Solving three-dimensional plasma physics problems at 1.71 GFLOPS

In an international effort to determine the magneto-hydrodynamic (MHD) stability of three-dimensional toroidal plasma equilibria, a team of researchers achieved a performance of 1.71 GFLOPS on a CRAY Y-MP computer system. "Plasma physics problems that used to take hours of computer time to solve in two dimensions now can be solved in 20 seconds in three dimensions," says David Anderson of Lawrence Livermore National Laboratory in Livermore, California. His teammates included Ralf Gruber and W. Anthony Cooper of Ecole Polytechnique Fédérale-Lausanne in Switzerland, and Ulrich Schwenn of the Max Planck Institute for Plasma Physics in Munich, West Germany.

The researchers achieved GFLOPS performance with the TERPSICHORE program by using the Autotasking feature of Cray Research's CFT77 Fortran compiler and the UNICOS scientific library routines for matrix multiply (MXM) and matrix inversion (MINV). Many thousands of integrals required to compute the Fourier representations of the problem could be solved as matrix-matrix products as embodied in the Cray routine MXM. Says Cooper, "Autotasking did such a good job that we abandoned our plans to multitask the code manually. We were able to achieve a parallelization factor of 7.35 on an eight-processor machine."



The magnetic field coil structure and the enclosed plasma in one sector of a Helias stellarator configuration, as modeled with the TERPSICHORE program. The stability analysis, which required 34 billion arithmetic operations, was performed in 20 seconds on a CRAY Y-MP system.

CUG meeting addresses user issues

The following report inaugurates a new feature of CRAY CHANNELS. Every other issue will now include a summary of events from the most recent meeting of the Cray User Group (CUG) along with comments from the group's president. CUG was formed by Cray system users in 1977 to provide a forum for the exchange of ideas related to Cray systems and their applications. The group holds two general meetings each year. In September, CUG held its second meeting of 1989 in Trondheim, Norway, and elected Mary Zosel of the Lawrence Livermore National Laboratory as its new president. Her remarks on the Trondheim meeting follow.

CUG has just finished another successful and enjoyable meeting, this time in Trondheim, Norway, sponsored jointly by the NTH University and SINTEF. Cray Research President Marcelo Gumucio addressed the split of the company and possible future directions for Cray Research. Cray Research's Project Everest also was described; the project is designed to implement Cray Research's plans to improve the installation, maintenance, and reliability of its system software. In addition, the Posix 1003.10 standards group was described, stressing the need for user sites to participate in the draft of a proposed standard for NQS. There was a report on Cray Research's Gigaflop award competition. Derek Robb, CRAY Y-MP product coordinator, shared some of Cray Research's future hardware plans with the CUG audience, including information about the CRAY Y-MP follow-on project.

Other activities at the meeting included several tutorials covering TCP/IP, performance, UNICOS security, Autotasking, and performance modeling; the NQS/Batch Birds of a Feather (BOF) had an active discussion of requirements for the NQS batch system and how it competes for system resources; the new BOF session for User Service Consultants was warmly received as sites traded information about the mechanics and problems of running a user consulting organization; the Management Special Interest Committee (SIC) was busy reviewing advance copies of CRI documentation for site service support plans; the CRAY-2 MIG provided time for sites to learn about future Cray Research CRAY-2 hardware and software support plans; several sites traded information about problems related to tuning the scheduling of their machines properly to fit their particular installations.

The annual CUG elections are held at the fall CUG meetings. As a result of the elections held in Trondheim, Mary Zosel, Lawrence Livermore National Laboratory, replaces Steven Niver, Boeing Computer Services, as president; Chris Lazou of the London Computer Centre is the new vice president; Gunter Georgi, Grumman, was re-elected secretary; and Claude Lecoeuvre, Commissariat a l'Energie Atomique, was elected director at large. Bjornar Pettersen, University of Trondheim, was named to the director seat vacated by the resignation of Marco Lanzarini, CINECA. I would like to thank outgoing members of the Board of Directors, Gary Jensen, NCAR, and Bob Price, Westinghouse, for their service to CUG. Also at the meeting, Jean Shuler, NMFECC, was named to replace Lazou as the CUG newsletter editor.

The CUG board of directors always has a number of organizational issues to address. We have been asked by Cray Research to name two members of the company's new Network Advisory Committee. Because of continuing interest of the CUG members in the problems of network storage to support supercomputers, the board has decided to form a new special interest committee to address mass storage management. The board also is looking for better ways to interact with third-party vendors at the CUG meetings.

We look forward to a continuation of these and other topics at the next CUG meeting, in Toronto, April 9-13, 1990. The theme of the Toronto meeting is Distributed Supercomputing. We hope to see representation from all Cray Research customer sites there.

Employees of organizations that are CUG members, employees of Cray Research, and guests may attend the semiannual CUG meetings. Following is the schedule of conferences for the next two years.
April 9-13, 1990, Toronto
October 1-5, 1990, Austin, Texas
April 14-18, 1991, London
September 22-27, 1991, Santa Fe, New Mexico

In a magnetic fusion reactor, a toroidal (doughnut-shaped) cloud of plasma, which is confined by magnetic fields, reaches temperatures of 100 million°C. If this plasma ring is not stably confined and is allowed to touch the wall of the metal containing chamber, the ring will live for only a few microseconds. The TERPSICHORE program enables researchers to study the design of fusion experiments for reactors in three dimensions to determine which designs can confine the plasma stably. Earlier MHD stability programs take much longer to run and allow study only in symmetrical devices.

Space shuttle analysis pared from 14 hours to 6 seconds

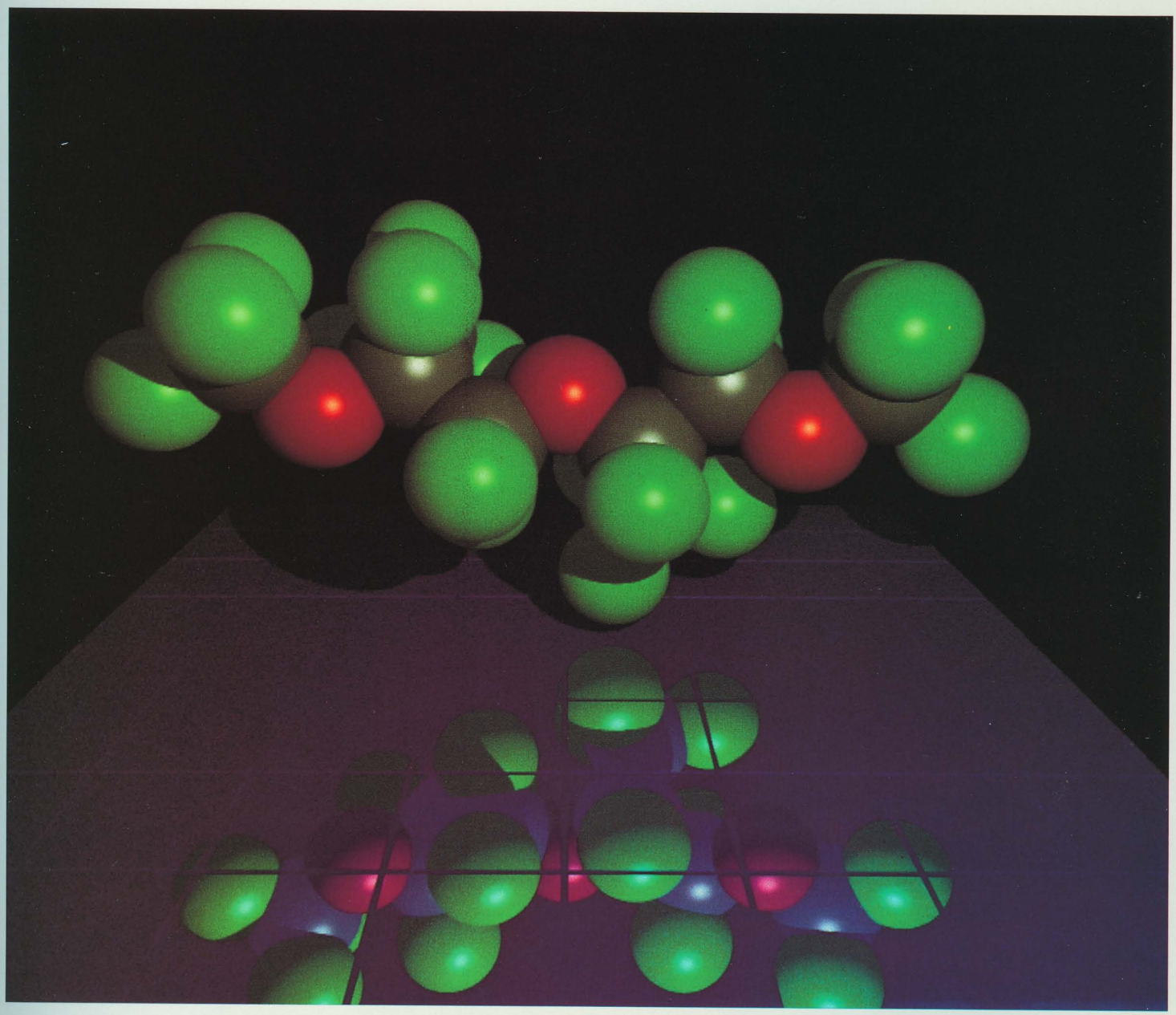
Olaf Storaasli of the NASA Langley Research Center and Duc Nguyen and Tarun Agarwal of Old Dominion University in Norfolk, Virginia, have developed a portable general-purpose algorithm that quickly solves simultaneous equations on computers with parallel and vector capabilities. When they applied the algorithm to the structural analysis of the Space Shuttle solid rocket booster and the solution of a three-dimensional cube, they achieved GFLOPS performance on a CRAY Y-MP system.

The Space Shuttle solid rocket booster originally was designed using a 1000-node (3000-equation) model that depicted the intersection of various rings. "After the Challenger accident, it was recognized that a more-detailed model was required," says Storaasli. A 54,870-equation model was developed to provide a more-detailed analysis of the lift configuration; this enabled scientists to examine the stresses created when the shuttle engines ignite and the boosters bend and return to their original positions.

"Since the model took over 14 hours to analyze on a VAX computer, it was moved to the CRAY-2 system, which reduced the problem to an hour and 6 minutes," says Storaasli. Although a vectorized version of the model ran in 14 minutes on the CRAY-2 system, the team was not yet satisfied with the performance, so they developed a new algorithm to run efficiently on parallel computer systems. When implemented on the CRAY-2 system, the problem solution using the new algorithm took 13 seconds, and on the CRAY Y-MP system it ran in 6 seconds at a rate of 1.34 GFLOPS. The simply generated three-dimensional cube, a common benchmark, exhibited even faster performance than the solid rocket booster.

"This work opens the door for complex structural calculations in a number of areas that haven't yet been considered for large-scale computations, such as nonlinear and composite structures analysis," says Storaasli. "A few years ago, only 1000-node models could be examined. Now we can solve much larger systems of 50,000 or 100,000 equations. Supercomputers make these problems tractable by reducing the time of analysis."

GALLERY



This model of the Krytox® polymer was computed on the CRAY X-MP/28 computer system at the Du Pont Experimental Station in Wilmington, Delaware. The Cray system took about 1100 seconds to create the 1280-by-1024 image using the OASIS ray-tracing program. Krytox is a high performance lubricant manufactured by Du Pont and used in applications such as high speed disk drives and vacuum pumps. The model was produced by Du Pont research chemist David Dixon and Cray Research senior systems analyst Patrick Capobianco.

CRAY CHANNELS welcomes Gallery submissions. Please send submissions to the address inside the front cover.