Ninth Annual Report
Interactive Graphics for Molecular Graphics System

TR83-02

February 1983

Frederick P. Brooks, Jr.
Michael Pique
James S. Lipscomb
Lynn F. TenEyck

The University of North Carolina at Chapel Hill
Department of Computer Science
CB#3175, Sitterson Hall
Chapel Hill, NC 27599-3175

This research was sponsored in part by NIH Grant #RR 00898-09.

B. Highlights - 1982-1983

Multiple Visualizations of Molecules Demonstrated

We and our scientific collaborators have developed and imported programs for producing many different geometric and parametric visualizations of molecular structure. This year we brought this work together in a focussed effort to apply all available techniques to a single molecule, Cu,Zn superoxide dismutase. In this way the various graphics techniques can be compared for their insight-inducing power and their costs. Moreover, combining all techniques on one molecule offers maximum insight into its structure and function.

Michael Pique produced a 17-minute videotape, "What Does a Protein Look Like?", that embodies some 40 visualizations of the SOD molecule. This together with a companion paper by F.P. Brooks, "Views of Unseen Worlds," was one of two invited keynote addresses at *Science Magazine's* Conference on Computers in Science in Chicago, in December. These works have been submitted for publication to SIGGRAPH'83, its proceedings (an issue of *Computer Graphics*,) and the *ACM SIGGRAPH Video Review*, a videotape journal. John Tainer and Elizabeth Getzoff of Duke University were Biochemist Consultants for the work.

Semiautomatic Density Map Interpretation

Tom Williams' system for computer-aided initial interpretation of electron density maps has been built and tested. Using a ridge line representation of the map, in which line segments connect local maxima, he has developed a method which saves a great deal of labor in the problem of going from the first electron density map to rough molecular coordinates. This reduces a step which normally takes weeks to a few days. Others have tried ridge line representations before, but not as an aid to human interaction. The ridge line representation appears to be the long-sought representation of electron density which permits computer graphics users to see large volumes of map without exceeding the capabilities of their displays, or, alternatively, to use standard displays instead of expensive specialized ones.

Real-Time Images of Space-Filling Models

M. Pique demonstrated real-time, joystick-controlled motion of space-filling (CPK) computer graphics molecular models of up to 50 atoms! While this is still short of the sizes needed for protein-enzyme research, it is a major technical achievement and offers promise for larger molecules in the future.

GRIP (mostly) Transported to VAX Computer

Our local GRIP-75 molecular graphics system, which has been extensively used for biochemical research, has been largely converted from an idiosyncratic hardware-software configuration that could not be readily duplicated elsewhere to a program in C, running under the UNIX operating system on the DEC VAX computer. Most of the system now works and it can be used on VAX-11/780 and 750 computers. Completion of this project will enable this useful software tool to be readily exported.

III. Narrative Description

A. Summary of Research Progress

1. Objectives and Operation

We have built, and operate as a service resource, an effective interactive computer resource for seeing, manually manipulating, and computationally modifying mathematical models of complex molecules. We believe that our present resource, called GRIP-75, has been shown to be as complete and useful as any in existence. One impressive measure of the power and utility of GRIP-75 is that at least seven of our clients have obtained their own graphics systems as a direct result of their successful work here.

Our resource has dual objectives. We are a service center providing powerful computer graphics facilities and expert computer assistance to chemists studying macromolecular structure. We are also computer scientists dedicated to advancing the art of interactive computation and interactive, three-dimensional graphics. The objectives are complementary. Our chemist clients provide the essential focus and a real, complex, and interesting driving problem for our computer science research; our computer science research in turn provides our clients with more powerful tools to improve their insight into very complex structures. Our overall objectives are:

1. To help our client biochemists obtain the most accurate structural information their data will support, with the most efficient use of their time, and
2. To advance the art of interactive graphics to the point that our clients are able easily to perceive the significant *chemical* information in their structures, at all levels of detail.

Our most valuable product is insight.

Seven years of experience with GRIP-75 have shown us a number of inadequacies. We are now building a second-generation version of this system, called GRIP-X, designed to be more comprehensive in the biochemical problems that can be attacked, more powerful in the mathematical tools available, more varied in the visualizations available for molecules and maps, smoother in the user interface, and constructed as product-quality, documented, exportable software.

Fundamental to our approach are the following principles:

- The GRIP system is designed to help chemists get results from their research, and its success is measured only by theirs.
- GRIP is designed to help the chemist visualize his molecules, his density maps, etc., so that he can use his knowledge to guide computational processes. That is, it is an aid to, not a surrogate for, human thinking and manipulation. Hence a strong emphasis is placed on human factors research and on human engineering of the system.
- GRIP is designed to serve many users, not one or two, so it includes an armory of alternative tools and techniques.
- GRIP is designed to interface smoothly with any batch computations its users must do, and to incorporate on-line facilities for all computations that can reasonably be done "while you wait."
- We as computer scientists are interested in GRIP as a test-bed for research in man-machine systems design, in man-machine interaction, and in the design of distributed computing systems.

A corollary of these objectives is that we are heavily dependent on observation of and feedback from real users attempting to solve real problems.

Our users are almost exclusively working on the structures of molecules of considerable biochemical interest: proteins and nucleic acids. We advance health-oriented biochemical research by enhancing the productivity of individual researchers through better tools. GRIP-75 was the trailblazing system in manipulating a molecular model to fit complex experimental data. Our clients have published over thirty papers in the biochemical literature containing results derived from the GRIP system. Features pioneered in GRIP are finding their way into other molecular graphics systems. The presence of the facility at UNC has been instrumental in maintaining a close collaboration with biochemists at Duke, who come over regularly to use our system. At least one of our new faculty members, Dr. R. Snodgrass, based his decision to come to UNC on the GRIP system, which demonstrated to him our ability to build large innovative systems.

Our facility consists of a VAX-11/780 computer, a Vector General 3303 vector display unit, an Evans & Sutherland PS 300 vector display unit, and a full complement of interactive input devices; and a PDP-11/45 computer with a Vector General VG3 vector display unit and a high-speed connection to the University of North Carolina Computation Center IBM 4341. Our present production system, GRIP-75, runs on the IBM 4341 and the PDP-11/45. The VAX-11/780 system is for the development of our next system.

The Computer Science Department obtained money from NSF for another VAX for departmental use. By pooling resources we have come up with a situation in which everybody wins. In exchange for some of the then-surplus capacity of our VAX we obtained, through the Computer Science Department and the Microelectronics Center of North Carolina, the use of 1.25 MB of memory (in addition to the 750 KB we bought) and a magnetic tape drive. We share the GRIP VAX with other graphics users in the department; in return we get the use of the Ikonas raster graphics system and the full resources of the department VAX — for example, the Versatec printer/plotter and a Hewlett-Packard four-pen plotter.

Our effective configuration on the VAX system is

- VAX-11/780 with UNIBUS but no MASSBUS;
- the UNIX† operating system (Fourth Berkeley version for the VAX)
- 2 MB memory
- 600 MB disk storage
- 800/1600 bpi dual density tape drive
- Vector General 3303 vector graphics display
- Evans & Sutherland Picture System 300 vector graphics display
- Ikonas RDS-3000 image processing and display system, with 1024 by 1024 pixels at 6 bits/pixel (or 512 by 512 at 24 bits/pixel), color map, two internal high speed processors, cross-bar switch for remapping pixel values, video digitizer, and write mask
- Summagraphics data tablet
- 16-channel (expandable to 32) analog to digital converter
- high speed parallel link to department VAX, which has hard copy plotters, printer, and a dial-up connection to the national "USENET" UNIX network

†UNIX is a Trademark of Bell Laboratories.

The usefulness of the USENET network of UNIX systems can hardly be overstated. News and mail can be exchanged with over 150 computer systems across the country, including systems at Columbia University, the University of California at San Francisco, the University of California at La Jolla, and the IBM UK Scientific Centre in Winchester, England. The systems mentioned specifically are all doing work in interactive computer graphics; several of them receive research support from the Division of Research Resources.

Another large gain from pooling resources with the department VAX has been reduction of maintenance costs. We can use each machine to verify faults in the other machine. Rapid identification of defective circuit boards translates directly to shorter and cheaper service calls.

2. Summary of GRIP-75 Use by year, 1975-1982

Table 1 summarizes the use of the GRIP-75 system for all purposes since we began demonstrations and productive operation on July 15, 1975. We have not tried to estimate the system time spent on development before the beginning of productive operation but know it to be many hundreds of hours.

Table I. GRIP-75 Use by years

	1975	1976	1977	1978	1979	1980	1981	1982	Totals
Production	329	581	781	1034	1108	556	349	130	4868
Demonstrations	12	50	161	137	43	57	76	52	588
Development	297	186	198	213	345	109	51	75	1474
Totals	638	817	1140	1384	1496	722	476	257	6930

Table II gives the use of the GRIP-75 system by year for each team of biochemists. These teams are identified by their principal investigators.

Table II. GRIP Production Time (Hours)

User	1975	1976	1977	1978	1979	1980	1981	1982	Total
Hermans	7	29	42	13	11	-	-	-	102
Kim	200	321	105	24	49	42	-	-	741
Richardson	83	79	91	188	384	297	169	57	1348
Lipscomb	12	-	-	102	-	-	-	-	114
Carter	27	21	-	4	5	16	-	-	73
Jensen		46	62	-	-	-	-	-	108
Tsernoglou		85	11	-	-	-	-	-	96
James			19	-	-	-	-	-	19
Low			55	191	65	-	153	-	464
Davies			85	92	-	-	-	-	177
Schiffer			35	28	26	28	-	15	104
Amma			163	35	-	-	-	-	198
Wright			41	-	96	-	-	-	137
Hendrickson			72	80	34	-	-	-	186
Schevitz				74	-	-	-	-	74
Love				94	-	-	-	-	94
Kartha				109	-	-	-	-	109
Taylor					89	45	17	44	178
Rich					78	-	-	-	78
Sarma					188	-	-	-	188
Olson					20	-	-	-	20
Bugg					39	-	-	-	39
Premilat					24	-	-	-	24
Amzel						98	-	-	98
Hardman						27	-	-	27
Eggleston						3	3	-	6
Roth							7	14	21
Total Hours	329	581	781	1034	1108	556	349	130	4868

Table III is a list of the biochemists whose research teams have used the GRIP system and their institutions. These are listed in the order of their first use of GRIP. For groups sending more than one biochemist to use our system, the principal investigator is given first and the names of his colleagues follow indented. The institutions of these colleagues are given only where they differ from the principal investigator's. In five cases —Sarma, Rich, Love, Sigler, and Amzel —the principal investigator has made little or no direct use of GRIP himself.

Table III. GRIP Users (1975-1982)

J. Hermans	University of North Carolina
J. E. McQueen	
D. R. Ferro	Istituto di Chimica delle Macromolecole
I. D. Kuntz	University of California School of Pharmacy, San Francisco
M. Vacatello	Institute of Chemistry, University of Naples, Italy
S. H. Kim	Duke University
J. L. Sussman	
R. W. Warrant	
S. R. Holbrook	
G. M. Church	
W. Shin	
C. J. Alden	
D. C. Richardson	Duke University
J. S. Richardson	
E. D. Getzoff	
J. A. Tainer	
D. McRee	
W. N. Lipscomb	Harvard University
J. L. Crawford	
C. W. Carter	University of North Carolina
R. A. Jones	
L. H. Jensen	University of Washington
K. Watenpaugh	
R. E. Stenkamp	Yale University
D. Tsernoglou	Wayne State University
G. A. Petsko	
M. James	University of Alberta
L. T. J. Delbaere	
G. Brayer	
B. W. Low	Columbia University
A. Sato	
M. Kimball	
S. Ginnell	
D. R. Davies	National Institutes of Health
M. C. Liu	
E. A. Padlan	
M. Schiffer	Argonne National Laboratory
E. L. Amma	University of South Carolina
R. L. Girling	
R. C. Paslay	
H. T. Wright	Princeton University
W. Hendrickson	Naval Research Laboratory

P. B. Sigler	University of Chicago
A. Podjarny	
R. W. Schevitz	
W. E. Love	Johns Hopkins University
W. E. Royer	
G. Kartha	Roswell Memorial Institute
H. C. Taylor	Berkeley Springs Research Consortium
A. Rich	Massachusetts Institute of Technology
N. Woo	
R. Sarma	SUNY, Stony Brook
A. Laudin	
A. Olson	National Resource for Computational Chemistry
C. Bugg	University of Alabama
R. Almassy	
J. Fontecilla	
S. Premilat	Universitéde Nancy, France
M. Amzel	Johns Hopkins University
K. Hardman	IBM Yorktown Heights Research Center
D. Eggleston	University of North Carolina
B. Roth	Burroughs-Wellcome
M. Corey	
L. Kuyper	
M. Connolly	Yale University

3. Changes in Resource Direction

The usage of GRIP-75 is rapidly declining. That is not unexpected. Computer graphics systems which do the same job as GRIP-75 are becoming fairly common. We believe that none of the other systems are quite as powerful as GRIP-75, but they are not that far behind.

In the past we have devoted much of our effort to client service and to research in how our clients made effective use of our system. (This is shown dramatically by the figures in Table I comparing production use and development use of GRIP-75 during the period 1977-1979.) We are now devoting nearly all of our efforts towards exploiting the things we have learned about human factors in the design and use of graphics systems, and to exploring advanced computer graphics techniques to be incorporated into our new system. These efforts are described much more fully in the next section. Once we have developed our new system we will again enter a major client service phase, this time by export of software.

Gifts for Resource Enhancement.

Our molecular graphics work has attracted much attention. This has resulted this year in substantial gift support for the Resource.

- Evans and Sutherland Computer Corporation donated a one-megabyte PS 300 system to UNC for the molecular graphics work. This high-performance vector- graphics instrument has a retail value of $89,000.
- Tektronix donated to our department two high performance color video monitors together worth $18,000. One of these is attached to the Ikonas raster-scan system and is used for molecular graphics.
- Vectrix gave us two low-speed 512-line color graphics displays together worth $10,000. One of these is configured so it can be used for molecular graphics.

- Intel Corporation gave us VAX memory worth $12,000. Half of this goes on the Resource's VAX.
- Burroughs-Wellcome gave us a cash gift of $13,000 to support an additional graduate research assistant on the GRIP project for 12 months.

4. Core Research - System Building

As the users of our GRIP-75 hardware-software have found it valuable, they have increasingly installed molecular graphics equipment of their own. Hence the work of our team has largely shifted from providing collaborative service to biochemists on our first-generation system to building our next-generation molecular graphics system, called GRIP-X. This core research dominated this year's activities.

4.1. Multiple Visualizations of Molecules

In GRIP-X, multiple ways of visualizing molecules, not just Kendrew (stick-figure) models will be important and possible. This year we imported and investigated all visualization techniques known to us, culminating in a research videotape by Michael Pique showing over 40 visualizations of Cu,Zn Superoxide Dismutase.

Views 5, 15, 16 were made by Tainer and Getzoff; the other pictures were done by Pique.

Views 22 and 24 are enclosed as samples. In view 22 the solvent-accessible surface has been colored to show the thermal activity of different parts of the molecule as determined crystallographically. The cool spot in the center shows that the copper is very rigidly bound. The cool spot to the left is where thw two parts of the symmetric dimer (of which only one is shown) attach rigidly back-to-back. View 24 shows a ribbon model of the backbone plus the copper and zinc. The backbone is colored by segment, showing clearly the β—barrel and the loops holding the metals.

Visualizations in "What Does a Protein Look Like?" Videotape

	Visualization	Program	Author
1	Coordinate Table		
2	Spinning backbone		
3	with side chains		
4	with hydrogens		
5	Ball-and-stick model	Ortep	Carroll Johnson (ORNL)
6	Peptide parallelograms	Ribvu	Pique
7	Shaded parallelograms	MOVIE.BYU	Hank Christiansen (BYU)
8	time-lapsed rotation	"	"
9	Smooth ribbon	Splined Ribvu	Pique
10	shaded rotating	MOVIE.BYU	Hank Christiansen (BYU)
11	Spherical model - color depth cue	Sphereshade	Pique
12	color by atom	"	"
13	with side chains	CPK	Thomas Porter (NIH DCRT)
14	fragmented spheres	Modified RAMS	Michael Connolly (Scripps)
15	Accessible surfaces - dots	MS	Michael Connolly (UCSF)
16	clipped section	"	"
17	varying probe radius	"	"
18	Accessible surfaces - smooth	RAMS	Michael Connolly (Scripps)
19	varying probe radius	"	"
20	color by species-invariance	"	"
21	by hydrophobic-hydrophilic	"	"
22	by thermal activity	"	"
23	by backbone segment	"	"
24	Ribbon, colored by backbone segment	MOVIE.BYU	Hank Christiansen (BYU)
25	Surface of back colored by segment	RAMS	Michael Connolly (Scripps)
26	Back showing β-sheet H-bonding	CPK	Thomas Porter (NIH DCRT)
27	Electrostatic field arrows	Amber	Paul Weiner (UCSF)
28	Animated electrostatics	Psanim	Pique
29	Electron density contours, 1 plane	GRIP-75	Britton
30	3 planes	GRIP-75	Britton
31	varying level	GRIP-75	Britton
32	Electron density ridge lines	Grinch	Williams
33	colored by interpretation	"	"
34	varying level	"	"
35	Close-contact parametric plot	MOVIE.BYU	Hank Christiansen (BYU)
36	Reciprocal-space Fourier map		Duncan McRee (Duke)
37	Ramachandran plot	GRIP-75	UNC team
38	dynamic scan	"	"
39	3-d rotation	"	"

4.2. Ridge-Line Density Representation and System for Initial Interpretation

Thomas V. Williams completed construction and testing of a prototype man-machine system for initial interpretation of density maps of proteins, a feat not heretofore accomplished by computer graphics. The breakthrough was the choice of Carroll Johnson's ridgeline representation for electron density instead of the conventional contour representation. This allows the biochemist to zoom back and forth between global and local views and to dynamically change the level of density visible.

By *initial interpretation* we mean estimating from an electron density map the positions of the mainchain atoms and perhaps the carbonyl and C-beta atoms. This method produces atomic coordinates for all of the atoms in each of the residues that have been located, using the amino acid sequence if it is available.

Williams's system is a man-machine system. Earlier attempts to interpret a map by a computer alone have not been very successful. Humans are very good at pattern recognition. They can bring together diverse sources of specialized information to determine whether some feature matches a pattern. They can balance a feature's match to a local pattern with the feature's global context to make an overall decision. These are things that computers do with much difficultly. Computers however are capable of accurate storage and recall of large amounts of information, pattern recognition for local, well-defined patterns, and numerical calculations.

The computer's task then is to remember the current state of interpretation and display it at all times. The ridge line graph and model are shown on a graphics screen with the edges color–coded by type. The different types of edges are molecular model, uninterpreted map, and map interpreted as each of the features mainchain, sidechain, carbonyl, and bridge. The bridge type designates disulfide bonds and hydrogen bonds. This encoding helps the user to see at a glance the portion of the mainchain traced so far, which edges attached to the mainchain have not been accounted for yet, and many other useful pieces of information about the current interpretation. The current display implementation does color line drawing on an Ikonas raster graphics system.

The user identifies features in the ridge lines and tells the computer which feature to associate with which part of the molecular model. The computer colors the features accordingly. The user adds edges as needed; the computer traces connected patterns. Extensive Undoing features permit easy backing out.

The system concept and command structure were set forth in detail in last year's report; we will not repeat. Since last year's report the system was tested on biochemist users and other maps. All the results are very encouraging.

Williams and Brian Van Duzee have this year devoted considerable effort to turning Williams's system from a working prototype into a robust product. We shall continue that in the coming year.

Ridge-Line Representation of the Electron Density Map

The most commonly used representation for electron density is contours drawn in sets of parallel planes. The contours show surfaces in the electron density function either at one or a small number of density values. The difficulty with using contours for initial fitting is that many lines are needed to accurately show the surface and a global overview is often necessary to resolve ambiguities. There is a limit to the number of lines that can be comprehended and interpreted at once. There is also a limit to the number that can be displayed flicker-free and dynamically rotated.

Ridge line representations of the electron density map were first suggested by Dr. Carroll K. Johnson of Oak Ridge National Laboratory. A ridge line representation shows the essential information about the location of high density and how it is connected, using many fewer lines than does a contour representation. Ridge lines, named by analogy to geographic terrain, form a network or graph connecting peaks (local maxima) through passes. If there was one and only one peak for each atom in the molecule then producing the ridge lines would completely solve the problem - each graph vertex would correspond

to an atom and each edge to a bond. For molecules the size of proteins this never occurs because of insufficient map resolution, phase errors, noise in the data collection process, atom mobility in the crystal, etc. The problem has been reduced to determining the correspondence between the ridge lines and the model. Some edges of the ridge line graph will lie along bonds, some edges are spurious and do not correspond to any part of the model, and some edges are missing, leaving bonds unaccounted for. The task then is to match the two graphs, the molecular model and the ridge lines.

System for Initial Interpretation

The use of ridge-line density map representations so reduces the number of lines to be displayed that Williams was able to build his prototype on the raster-scan Ikonas instead of the vector-drawing Vector Graphics 3303. This gave him color, which radically simplified his user interface.

Two corollaries emerge for GRIP-X. First, we need color on our vector systems, so that we can show quite complex molecule density maps and dot surfaces. Hence we are cutting back everywhere to squeeze such a color tube into the budget.

Second, Williams was able to deal with real maps using raster scan technology. This is radically cheaper than vector-drawing technology because many of the components are mass-produced for television. By using ridge-line representations of density for fitting as well as for interpretation, we hope to be able to build a full-function version of GRIP-X that will run on a color workstation costing under $35,000, such as the SUN.

Our experience with Williams's prototype suggests three exciting possibilities to be explored--none has been demonstrated:

- Jane Richardson's experience suggests that the overview zoom and variable contour level features of the system may make it possible to interpret lower-resolution or noisier maps than heretofore could be interpreted at all.
- The high quality of some of the initial interpretations suggests that it may be possible to skip entirely the usual manual fitting step for proteins, and go directly to algorithmic refinement by computer.
- The reduced picture complexity may permit us to move molecular graphics from high-cost instruments into the mainstream of computer graphics technology, permitting most protein structural chemists to have systems in their own labs.

Hence our plate is full!

4.3. Space-Filling (CPK) Models

A major goal for GRIP-X is to allow real-time moving and docking of CPK (solid spherical) models of molecules. We this year demonstrated real-time motion of molecules of up to 50 intersecting atomic spheres!

Just as contours are not the only nor necessarily the best representations of electron density, so stick-figure (Kendrew) models are not the only representation for molecules. Biochemists routinely use plastic space-filling models composed of spherical atoms (CPK models) as well, particularly when investigating packing and docking.

With computer graphics, it has heretofore been possible to use stick-figure models with dynamic motion or colored and shaded spherical models for static

viewing only. The calculations for spherical models could not be done in the 1/ 30 to 1/ 12 second necessary for dynamic motion.

On GRIP-X we want to achieve dynamic motion for space-filling models, and we think we can within the next few years.

This year Dr. Barry Zeeberg of UNC used the facility to continue development of a technique suggested by Pique in which each atom is represented by a circular disk parallel to the viewing screen. The disks are painted to look like shaded spheres. Of course, they aren't; at any given time one disk may partially overlap another, but its entire surface is in front of or behind the other's; there are no *intersections* visible. As the molecule is rotated, disks jump discontinuously in front of or behind each other. Nevertheless, it is a more rapidly computed approximation to a space-filling model.

Pique explored ways to use the Ikonas display's very fast internal processor to construct realistic pictures of properly intersecting spheres capable of dynamic movement. In January he demonstrated dynamic movement of molecules of up to 50 atoms. While still an order of magnitude slow for proteins, this is an exciting and encouraging achievement.

4.4. GRIP-75 Conversion

In spring, 1982, a COMP 145 student team consisting of J. Draper, M. Miller, H. Taylor, and M. Wright undertook to convert parts of GRIP-75 from PL/I to C, from the IBM System/360-75 to the DEC VAX-11/780, from the OS/MVT operating system to the UNIX operating system. It worked much better than we had expected. Substantial parts of GRIP-75 ran correctly on the VAX by the end of the spring term.

That team used the *pltoc* PL/I to C translation aid built by the spring 1981 COMP 145 student team of C. Burns, A. McDonald III, M. Moore, and R. Smith. They used a powerful UNIX tool called YACC, which they fed the PL/I syntax tables in the Vienna Definition Language, straight out of the report of the IBM Vienna Laboratory.

During the year Pique, aided by Rich Hammer, have continued this work. It now appears that we may be able to run GRIP-75 entirely from the VAX within another year. Moreover, this conversion permits us to take GRIP-75 code bodily into GRIP-X as that seems wise.

4.5. Collaborative GRIP-X Design: Database and Adjunct Language

We continued our collaboration with the IBM UK Scientific Centre (UKSC) in Winchester, England, who are building a GRIP-X-type system in PL/I for an IBM host and a Vector General display. Brooks visited there in May; Dr. William Wright and Dr. Geoffrey Robinson visited here in September; Dr. Wright will be here again in March. We confer weekly by a conference call between the two teams. We finally have reliable electronic mail via the UNIX UUCP-net connected to the IBM VNET. We have spent days discussing overall system design, Wright's *agent* concept, and how the system partitions.

Database. The UKSC team came up with a simply superb idea--to incorporate an existing relational database system within GRIP-X instead of building elaborate idiosyncratic data management facilities. We see it working this way: When a new visualization (or *view* in the database sense) is called for, a query goes to the database system, which produces the relations requested. These are then compilatively translated into an intermediate form from which pictures can be rapidly generated and in which manipulations can be rapidly effected. Upon Save or Quit commands, new data are posted back into the database.

Stephen Todd and Andrew Morffew of the UKSC team did tests with a system UKSC had built earlier, the Peterlee Relational Test Vehicle. These tests showed that even without any tuning, responses to demanding queries could be made in 15-30 seconds, which seems acceptable considering the low frequencies of queries we expect.

The UNC team imported and began investigation of Stonebreaker's INGRES, a Unix-based relational database system. Because the user interface seems better, we dropped INGRES and have imported Wasserman's TROLL, another UNIX–based relational system. We plan to investigate it more in the coming year.

Adjunct Language One of the powerful attributes of GRIP-75 is its implementation and use of a geometric manipulation language between the menu language and the PL/I (or C) level. This has enabled us to build new menu commands for users very quickly, adapting the system to user needs.

Pique, Lisa Weber, and other team members have pursued various issues in the design of an adjunct language for GRIP-X. We do not yet have a design, but we know some important components.

4.6. Advanced Graphics Techniques

At all times we keep exploratory efforts going on advanced ideas and technology. Often these are done at no cost to the GRIP project; they are carried on by other funded projects or by student team efforts in our COMP 145 course--Software Engineering Laboratory. The things that didn't work or don't work yet are enumerated in the next section. Some of the ones that did are discussed in the section above. Others are sketched here.

SUN Terminals. Graphics workstations built on the Motorola MC68000 processor offer high performance at low cost. We believe they will be a significant carrier for GRIP-X.

We have selected the workstations of the Stanford University Network (SUN) design for our experiments. This year Brian Van Duzee and Doug Schiff attempted to transport Tom Williams's system for the VAX to a department-owned SUN workstation built by Cadlinc. Due to operating system inadequacies in the Cadlinc Unix System, we suspended this effort until our project's our SUN workstation built by SUN Microsystems Incorporated arrives. In this way we will debug the software on the machine we plan to use, one with a color system. It is due for installation next week.

Pixel-Planes. Prof. Henry Fuchs, a department faculty member, has invented a custom VLSI chip that should radically speed up processing of polygon definition, hidden surface calculations, and polygon color interpolation. A second prototype version is now running; a third is about to be fabricated.

Brooks studied how this might be applied to molecular graphics. The chip can be programmed to evaluate expressions of the form $Ax + By + C$ at all pixels concurrently. Brooks devised a program for the chip that allows multiple circle definitions to be evaluated in about the same time each as one edge of a polygon. The surprising result is that circle evaluation can be done at all with a linear expression evaluator.

Brooks also devised an algorithm by which Pixel-Planes can be programmed to do "bump-detection," the instant detection of an attempt to move one molecular fragment into a volume element occupied by another. This promises to be useful for drug-design studies.

Varifocal Mirror. Traub of MITRE Corporation in 1967 patented a true 3-D display made by the synchronous plotting of points on a vibrating curved mirror.

Hobgood and Brooks built one of these in 1970 and studied user perceptions of it. Fuchs, Prof. S.M. Pizer, and Brooks are investigators in another project to develop this technology.

This year that team showed for the first time that protein molecule stick figures can be rotated in real time while being thus displayed.

Head-Mounted Display. In a separate project, Fuchs and Gary Bishop are building in 1983 technology a reincarnation of Ivan Sutherland's 1960's head-mounted display, which never quite worked well enough for productive use.

The idea is that the user views a virtual world in which he is embedded, rather than looking at one through the porthole offered by a normal graphics screen. Separate displays are offered to the left and right eyes through half-silvered mirrors that superimpose virtual objects on the real world. As the position and orientation of the head changes, the scene is updated in real time.

Bishop has running a slow-motion prototype in which the user sees a protein molecule filling the space, "life-size," as it were. The hard problem is tracking the position and orientation of the head. Fuchs and Bishop have a new scheme for that now under implementations. Brooks and Pique are advisors to that project.

Raster Analytic Molecular Surfaces. Dr. Michael Connolly of Scripp's Clinic Research Institute in La Jolla visited our facility for 10 days in August 1982 to collaborate with Pique in developing Connolly's shaded-surface display of molecular surfaces. Like his well-known dot surfaces, these show the molecular surface accessible to a probe sphere of a given radius. Unlike the dots, or tiled surfaces derived from the dots, the surfaces generated by his "RAMS" program are continuous, analytically smooth, and cover the molecule with no gaps or cracks. The program runs on our VAX computer and generates images for the Ikonas color raster display. Each view of a typical protein requires about two hours to calculate. About 60 such images were included in Pique's videotape, described above.

4.7. Busts

We also tried some things that either didn't work or else don't work yet. Some of these we will come back to later.

Molecular syntax. We investigated the codification into rigorous production-rule form of Jane Richardson's informal rules for what topology and shapes do and don't occur in known molecular conformations. Too hard.

Crystallographic computation package. We tried packaging and transporting into our system a set of high-performance Fortran programs for crystallographic computations previously developed by Lynn TenEyck. It turned out that some of the component programs had bugs too difficult for our COMP 145 student team to solve during the semester, and Lynn hasn't had time to tackle them himself.

Sinclair color display for VG 3303. We provided a test bed and collaboration for Larry Sinclair to come test his newly-developed color tube for the VG 3303. The tube worked and drew color pictures, but the power supplies for the deflection circuits were grossly inadequate, so the pictures shrank. Sinclair went back to California to work some more. We did not buy his system as budgeted. Instead we bought a color front end for the new SUN workstation.

5. Collaborative Research

5.1. Richardsons, Tainer, Getzoff - Superoxide Dismutase.

Work this year culminated in Ph.D. dissertations by Tainer and Getzoff showing the detailed structure and mechanism of catalytic action of Bovine Cu,Zn Superoxide Dismutase.

5.2. Richardsons - Studies on Interpretation of Density Maps

Drs. D. and J. Richardson of Duke University are continuing a research project to study the effect of resolution on the accuracy of the interpretation of electron density maps. They believe that in maps with poorer than 3 A resolution certain types of misinterpretations become common. They are attempting to confirm this in controlled studies using maps calculated to different resolutions of known structures. There is surprisingly little systematic knowledge of the effect of resolution on the accuracy of a structure, partly because it is very difficult to separate the effects of resolution and phase errors. This study should give insight into the nature of low resolution maps, and could act as a valuable guideline for crystallographers deciding how much work to invest in data collection before the attempt to interpret their maps.

5.3. Hermans, Carson, Watenpaugh, Van Gunsteren - Molecular Dynamics

About half of the CPU cycles of the VAX-11/780, but none of GRIP-75, was used extensively by this team for developing:

- a package of software tools for studying the dynamics of protein and solvent molecules in crystals,
- a new macromolecule refinement procedure with energy minimization and least-squares crystallographic refinement,
- an excluded-volume theory of polymer-protein statistics tested by Monte Carlo Simulations.

Dynamics of protein and solvent molecules in crystals

Jan Hermans, Mike Carson, W. van Gunsteren - Biochemistry, University of North Carolina.

After extensive program development, we are presently performing simulations of some scientific interest.

A number of programs have been brought into a unified format:

PREDATOR- prepares the data describing a system to be simulated in a form usable by the other programs, from easy-to-understand input.

CEDAR- performs molecular dynamics or energy minimization; the program allows for presence of any number of molecules in the system, has special features for rapid calculation of water-water interactions, and maintains crystal symmetry.

MONTE- performs Monte Carlo calculations on the same sort of systems, moving and rotating water molecules,

ANALYSIS- Various features within Cedar, and several stand alone programs perform: Calculation of structural parameters, calculations of average and rms values, calculation of simulated density maps, calculation of time correlation functions.

GRAPHICS- A program was written that will perform playback of short time sequences of a molecular dynamics run, as images of stick models with dotted hydrogen bonds, on the Evans and Sutherland PS 300. Another program shows a stick model with thermal parallelopipeds at each atom location.

Simulations in Progress

Simulation of the Alpha Helix. Four simulations are being performed: of an 11-residue molecule, (alanine)*11, both in vacuo and in a box with water molecules at liquid water density, and of an infinite polyalanine molecule with 11-residue repeat (=3 helical turns), again both in vacuo and in liquid water. Pending a thorough analysis, we can already say that the structures vary more in water than in vacuo; in fact, the 11-residue molecules becomes so "messy" during short molecular dynamics runs that we believe they are undergoing unfolding to the random-coil conformation, or rather, to different members of the enormously large family of conformations that together constitute "the" random-coil.

Building Guessed Structures. At Charlie Bugg's request, we have simulated static conformations of two proteins of unknown structure, according to his hypotheses. Both structures were built by analogy with existing x-ray crystallographic structures of other proteins. For one of the proteins (a scorpion toxin) this was an obvious analogy between proteins of similar function but from different species. The other protein, or rather, peptide, "apamin" from bee venom, was built in analogy with a small part of the scorpion neurotoxin. The resultant structure showed a number of characteristics that agreed with recent conclusions based on 2-D NMR measurements that were unknown to us when the simulation was performed.

MACREF, a macromolecule refinement procedure with energy minimization and least-squares crystallographic refinement. Jan Hermans, Mike Carson, UNC; K.D. Watenpaugh, UW, Seattle.

The program contains both PREDATOR and CEDAR. Both programs have been extensively edited to be in conformance with the prescribed format of the Xtal-80 crystallographic programs. Mike Carson spend some time in Seattle learning to do this from K.D. Watenpaugh, and Jan Hermans is about to spend a week in Seattle to assess performance of the program in refinement. The work will be reported at the West-Coast protein crystallography conference in March 1983.

Excluded-volume theory of polymer-protein interactions based on polymer chain statistics.

An analytical theory has been developed and has been thoroughly compared with results of Monte Carlo simulations of exclusion of spheres by random segmented chains (and vice versa). The agreement is essentially perfect. Results of the theory also agree well with available experimental thermodynamic data, of which some were obtained in our laboratory, and others elsewhere.

One paper has appeared, and one has been accepted. A technical paper on the simulation method remains to be written.

J. Hermans, "Excluded-volume theory of polymer-protein interactions based on polymer chain statistics", *Journal Chemical Physics*, 77, 2193-2203 (1982).

D. Knoll and J. Hermans, "Polymer-protein interactions: Comparison of experiment and excluded-volume theory," *Journal Biological Chemistry*,

accepted for publication.

These papers bear the following acknowledgment: "Computations were done with support from the National Institutes of Health, Division of Research Resources (grant RR-00898)."

This work has been/is being supported by grants from NSF (PCM81-12234) and NIH (HL-26309). NSF supports molecular dynamics and simulations in general, NIH supports, amongst others, a study of polymer-protein interactions.

5.4. Roth, Corey, Kuyper - Interactions of Drugs with Proteins

We are continuing a collaboration with scientists at the Burroughs-Wellcome Company in Research Triangle Park, North Carolina, to develop improved methods for studying drug interactions with proteins and the evaluation of compounds for possible biological activity.

Initially we have aided in studies by Dr. Barbara Roth of the binding of trimethoprim to the enzyme dihydrofolate reductase. So far this work has been a rather basic docking exercise with a small number of degrees of conformational freedom in the trimethoprim. We are using this simple initial study as a prototype to determine what tools will be appropriate for this task.

Dr. Mike Corey of Burroughs-Wellcome uses the PROPHET system to generate and maintain files of chemically related compounds as part of the drug design process. We have been discussing Dr. Corey's research problems with him to determine just which parts of his work are most suitable for automation and graphical aids. Doris Knecht is our liaison person with the Burroughs-Wellcome team. As an interim measure, a COMP 145 team has developed an interface between our VAX and PROPHET which enables a PROPHET user to use our high performance three-dimensional graphics viewing system instead of Tektronix storage tube graphics. Scott Hennes continued this work during 1982.

5.5. Wolfenden - Molecular surface of SGPA

Dr. R. V. Wolfenden of the UNC Department of Biochemistry is using the facility to aid his study of the molecular surface of the SGPA enzyme, whose structure has been determined by Michael James and colleagues at the University of Alberta. SGPA (*Streptomyces griseus* protease A) is a proteolytic enzyme with sequence homology to the trypsin family of serine proteases only in the immediate vicinity of the active site. Our team used the published coordinates of SGPA as input to Michael Connolly's "MS" molecular dot-surface program, and displayed the surfaces with superimposed stick-figure molecular models on the Vector General 3303 and Picture System 300. Dr. Wolfenden is writing James to ask for coordinates for an SPGA substrate; when those arrive we will calcuate the substrate's independent dot surface and see how well it fits the enzyme's. Having a color monitor for either the VG or the PS would greatly aid this work. Wolfenden is also using the numerical solvent-accessibility estimates from MS to guide protein-folding predictions.

5.6. Topal - Sequence-dependent Effects of DNA Methylation

Dr. Michael D. Topal of the UNC Department of Pathology has used the Vector General 3303 and Picture System 300 displays to view molecular dot surfaces computed by Connolly's MS program. He is trying to predict the effect on DNA conformation of adding methyl groups to particular rings in certain base sequences. His work needed molecule-editing powers beyond what we can now provide, so he is currently trying to open a Prophet account. We expect to work

with him again this spring.

5.7. Schiffer - 5 A resolution structure of Immunoglobin LOC

Dr. Marianne Schiffer of Argonne National Laboratory visited UNC for four days to try to learn the gross conformation of the Immunoglobin LOC. Starting with the refined structure of a similar molecule previously fit on GRIP-75, Bence-Jones immunoglobin MCG, she used GRIP-75 and Williams's ridge line system to try to identify the constant and variable domains in the new molecule's low resolution map. The work was difficult, as neither system was designed for such low resolution maps. She was able to locate the constant domains confidently, and the variable domains tentatively, but decided to go home and work toward a better map. She now has a 3.5 A map, and is scheduled to visit us again the second week of March, 1983.

6. Training

One visiting scholar, Dr. Thomas Hern of Bowling Green State University, and five new graduate research assistants joined the project this year. The training of new research assistants was a major project for the Fall term; Lisa Weber did this training. Several of these will be with us for three or more years.

We give our clients sufficient training to use the system when they arrive, but do not train them in the operation of the underlying computer systems. Instead we provide them with trained assistants as "shepherds" to take care of that aspect of their use of the system. We find that it takes one to two days for a client to become comfortable with the graphics and structure manipulation features of GRIP-75, and the remainder of the visit can be devoted to productive work.

B. Highlights 1982-83

Multiple Visualizations of Molecules Demonstrated

We and our scientific collaborators have developed and imported programs for producing many different geometric and parametric visualizations of molecular structure. This year we brought this work together in a focussed effort to apply all available techniques to a single molecule, Cu,Zn superoxide dismutase. In this way the various graphics techniques can be compared for their insight-inducing power and their costs. Moreover, combining all techniques on one molecule offers maximum insight into its structure and function.

Michael Pique produced a 17-minute videotape, "What Does a Protein Look Like?", that embodies some 40 visualizations of the SOD molecule. This together with a companion paper by F.P. Brooks, "Views of Unseen Worlds," was one of two invited keynote addresses at *Science Magazine's* Conference on Computers in Science in Chicago, in December. These works have been submitted for publication to SIGGRAPH'83, its proceedings (an issue of *Computer Graphics*,) and the *ACM SIGGRAPH Video Review*, a videotape journal. John Tainer and Elizabeth Getzoff of Duke University were Biochemist Consultants for the work.

Semiautomatic Density Map Interpretation

Tom Williams' system for computer-aided initial interpretation of electron density maps has been built and tested. Using a ridge line representation of the map, in which line segments connect local maxima, he has developed a method which saves a great deal of labor in the problem of going from the first electron density map to rough molecular coordinates. This reduces a step which normally takes weeks to a few days. Others have tried ridge line representations before, but not as an aid to human interaction. The ridge line representation appears to be the long-sought representation of electron density which permits computer graphics users to see large volumes of map without exceeding the capabilities of their displays, or, alternatively, to use standard displays instead of expensive specialized ones.

Real-Time Images of Space-Filling Models

M. Pique demonstrated real-time, joystick-controlled motion of space-filling (CPK) computer graphics molecular models of up to 50 atoms! While this is still short of the sizes needed for protein-enzyme research, it is a major technical achievement and offers promise for larger molecules in the future.

GRIP (mostly) Transported to VAX Computer

Our local GRIP-75 molecular graphics system, which has been extensively used for biochemical research, has been largely converted from an idiosyncratic hardware-software configuration that could not be readily duplicated elsewhere to a program in C, running under the UNIX operating system on the DEC VAX computer. Most of the system now works and it can be used on VAX-11/780 and 750 computers. Completion of this project will enable this useful software tool to be readily exported.

C. Administrative Changes

The major administrative change is that Dr. Ten Eyck left in June to return to Oregon. His position was advertised in *Science* and *Nature*. Several candidates were interviewed; none were entirely suitable. We are leaving this position unfilled for year 10 in order to make ends meet and buy an essential piece of equipment not planned in 1978.

D. Resource Advisory Committee and Allocation of Resources

Table IV lists the members of our Advisory Committee.

We currently offer the facility, and such help as we can give, free of charge to any chemist:

- who has a scientifically interesting problem,
- whose work is at a stage where our facility might be useful,
- who is willing to commit his time, travel money, and effort to a serious use of the facility, and
- who is willing to give us written and oral feedback from his experience.

Table IV: Advisory Committee Members

Name	Degree	Title	Department	Institution
F. P. Brooks	Ph.D.	Kenan Professor & Chairman	Computer Science	UNC-CH
J. Hermans	Ph.D.	Professor	Biochemistry	UNC-CH
D. Richardson	Ph.D.	Professor	Biochemistry	Duke U.
J. Richardson	M.S.	Assistant Professor	Anatomy	Duke U.
W. V. Wright	Ph.D.	Senior Systems Architect	UK Scientific Centre	IBM England

E. Dissemination of Information

The availability of GRIP-75 is widely known among crystallographers. We publicize the facility by announcements and notices at scientific meetings, by demonstrations to all interested parties, and by word of mouth. Demonstrations are perhaps the most effective means.

Announcements of Availability

During the 1982-1983 grant period the availability of GRIP-75 was announced at the Washington, D. C. meeting of the American Crystallographic Association in March, 1982. In addition GRIP-75 was included in the list of molecular graphics facilities published by the Brookhaven Data Bank, and of course is in the list of Biotechnology Research Resources published by DRR. These announcements should reach all of the scientists who might be interested in our facility. Our experience indicates that most of our clients hear about our facility from other crystallographers, then look up our telephone number in the DRR Biotechnology Research Resources list.

Michael Pique maintains an internationally recognized census of (over 60) molecular graphics installations. At SIGGRAPH'82, he organized a special informal session on molecular graphics, at which our installation was described, among others.

Demonstrations

Some 52 hours of 1982 GRIP-75 resource time were used in demonstrations, plus many more hours of VAX time. During 1982 we demonstrated GRIP to the following biochemists and computer scientists.

Mr. H. Stover	Genisco Corp.
Dr. C. D. Barry	Washington Univ.
Dr. Cyrus Levinthal	Columbia Univ.
Dr. W. Gilbert	Mass. Inst. of Technology
Mr. J. McAllister	Washington Univ.
Dr. D. Rohrer	Medical Research Foundation
Dr. James W. Fawcett	General Electric Electronics Lab
Dr. Noble R. Powell	General Electric Electronics Lab
Dr. Richard Economy	General Electric Simulation & Control Systems
Mr. Terry R. Fowler	General Electric Calma Company
Mr. Geoffrey Robinson	IBM UK Scientific Centre
Dr. Thomas Hern	Bowling Green State University
Dr. Harvey Cragon	Texas Instruments
Dr. Richard Zayre	Stanford University
Dr. Charles Johnson	Univ of North Carolina Chemistry Dept.
Dr. Thomas Pierce	Rohm & Haas Pharmaceuticals
Dr. John Ritz	Rohm & Haas Pharmaceuticals
Mr. Joseph Gilbert	Rohm & Haas Pharmaceuticals
Dr. Harold Almond	McNeil Pharmaceuticals
Mr. Steven Satterfield	US Naval Academy
Dr. Steven Anderson	Yale University
Mr. Eric Goldberg	Tektronix Corp.
Dr. Howard Struble	Duke University
Mr. Richard Feldmann	National Insts of Health
Dr. W. van Gunsteren	Univ of Groningen

Dr. Krauss	Frederick Cancer Resc. Center
Dr. Elmore Millner	Union Carbide
Ms. Rebecca Spitz	Evans & Sutherland Corp.
Mr. Tim Bleakley	Evans & Sutherland Corp.
Dr. F. Richards Adrion	National Science Foundation
Dr. Chien	National Science Foundation
Dr. W. Wilner	American Bell Labs
Dr. Andries van Dam	Brown University
Dr. Vivian Alured	Univ of Colorado
Dr. Gene Faulcon	US Environmental Prot. Ag.
Dr. Lowell Harris	Mayo Medical School
Dr. Jon Camp	Mayo Medical School
Mr. Richard Brady	Sidwell Friends School
Dr. Robert St.John	Ciba-Geigy Pharmaceuticals