First Annual Report
Molecular Graphics Project

TR75-03

March 1975

Frederick P. Brooks, Jr.
Jon Bentley
Edward Britton
James Foley
Jan Hermans
James Lipscomb
James Robb
Joel Sussmann
William Wright

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

NATIONAL INSTITUTES OF HEALTH
DIVISION OF RESEARCH RESOURCES
BIOTECHNOLOGY RESOURCES BRANCH

SECTION I - RESOURCE IDENTIFICATION

Report Period:		Grant No.
From: July 1, 1974 To: June 30, 1975		1-P07-RR00898-01
		Date of Report Preparation
		February, 1975

Name of Resource	Resource Address	Resource Telephone No.
Molecular Graphics System	273 Phillips Hall, U.N.C. Chapel Hill, N. C. 27514	-

Principal Investigator	Title	Academic Dept.
Dr. F. P. Brooks, Jr.	Professor & Chairman	Computer Science
Dr. J. D. Foley (Acting: 1/75 - 7/75)	Assistant Professor	Computer Science

Grantee Institution	Type of Institution	Investigator's Telephone No.
University of North Carolina at Chapel Hill	State University	919/933-7339 (Foley)

Name of Institution's Biotechnology Resource Advisory Committee:
Scientific Advisory Committee

Membership of Biotechnology Resource Advisory Committee:
(* have reviewed this report) (** Committee Chairman)

Name	Title	Department	Institution
Frederick P. Brooks, Jr.**	Professor & Chairman	Computer Science	UNC
Ernest L. Eliel *	Professor	Chemistry	UNC
Jan Hermans *	Professor	Biochemistry	UNC
Sung Hou Kim *	Associate Professor	Biochemistry	Duke
Edward Perl *	Professor & Chairman	Physiology	UNC
David Richardson *	Assistant Professor	Biochemistry	Duke

Typed Name & Title of Principal Investigator	Signature	Date
James D. Foley, Assistant Professor		

Typed Name & Title of Grantee Institution fficial	Signature	Date
L. Felix Joyner, Vice President - Finance University of North Carolina	L. Felix Joyner	3/5/75

delete for new area.

II.A. DESCRIPTION OF RESOURCE OPERATIONS AND PROGRESS.

The purpose of our research is to make the power of interactive computer graphics available to biochemists and crystallographers seeking to understand the conformation of complex molecules such as proteins and nucleic acids. This is an exceedingly difficult problem, and is currently addressed, in large models, usually with the aid of electron density maps and the Richards Box. The problems with traditional approaches, especially for large molecules, are numerous. We are developing two major interactive computer programs, the one for conformational manipulation and modification based on idealized local conformations and internal energy bonding considerations, the second as an Electronic Richards Box. Each system is being developed by working with a biochemist-user; Dr. Jan Hermans and colleagues in the former case, Dr. Sung-Hou Kim and colleagues in the latter. As the systems are shaken down and become solidly operational, we will actively seek additional users. This will probably happen in Fall of 1975.

Our work is currently done with the UNC Department of Computer Science's graphics facility, which includes a Vector General display, a PDP-11/45 minicomputer dedicated to graphics use, and an IBM 360/75 computer for non dedicated graphics use. The resource will soon be expanded with the delivery of a Tektronix 4014 display terminal, to be located at Duke in Dr. Kim's lab. It will ultimately be used to access the interactive programs mentioned above, and can of course be used in other ways as well.

To support the development of the interactive graphics application programs, we are building more powerful programming tools as required by the applications, and are maintaining and developing our existing programming tools.

II.A.1. Scientific Structure and Work in Progress.

We have organized by mission; there are five active groups working on as many projects. We consider all projects as Core Research and Development, except for some of Dr. Hermans' work, which is collaborative in nature.

The first project, called GRIP (Graphics Representation of Interacting Proteins) is directed by Dr. William V. Wright. Dr. Wright's services are made available to the project through a joint study agreement between UNC and his employer, IBM. Dr. Wright is assisted by Mr. Michael Pique.

The initial tester of GRIP will be Dr. Jan Hermans. He has been continually interacting with the development team, to assure that GRIP meets his needs. The purpose of GRIP is to allow a biochemist to study and change a molecular model, in an attempt to bring it into closer correspondence with observed or hypothesized characteristics. Briefly, GRIP allows a large molecule to be viewed, in whole or in part (optionally in stereo), dynamically rotated, and modified. A general purpose refinement algorithm, discussed below, is being integrated into the system.

A second team, directed by Dr. Hermans and including Drs. Dino Ferro, John McQueen, Shirley Wei, and Mr. Jon Bentley, are further developing the conformational refinement algorithm initially developed by Hermans and McQueen [1]. A program, called REFINE, uses the algorithm to move a trial molecular conformation into closer agreement with constraints on atom position, geometry (bond lengths, bond angles, and dihedral angles), and nonbonded energy.

In addition to developing REFINE, Dr. Hermans' group has been helping to integrate it into GRIP, and to demonstrate its use as a tool in the interactive manipulation of large structures. They are also engaged in collaboration with several groups outside UNC who are using REFINE in their own work. Mr. Bentley is doing theoretical work on developing an algorithm which may make REFINE work faster.

A third effort, conducted by Mr. Edward Britton and supervised by Dr. Brooks, has as its goal the development of an "Electronic Richards Box". This system, called GRAB, is an extension to GRIP which will allow a user to interactively build and modify a molecular model to conform with an electron density map. The first user of GRAB will be Dr. Sung-Hou Kim; he has therefore served as adviser in the development of GRAB.

At Duke, Drs. Kim and Sussman are formulating plans for the use of their Tektronix 4014 direct-view storage tube display terminal, delivery of which is expected in March or April. Initially they will modify existing FORTRAN batch programs to use the display. Then, when the display-independent graphics package we are developing is operational, GRIP and GRAB will be made usable from the 4014,

although the interactions will certainly be less rapid and less satisfactory than with the Vector General refresh display, and real-time manipulation will not be available.

Finally, we are undertaking to improve and further develop the hardware and software tools available for use by the developers of the interactive graphics applications programs. This facility development work is directed by Drs. Foley and Wallace, with assistance in some areas from Mr. Robb. Research assistants are Ms. Strode, and Messers. Hogan, Lipscomb, Babich, and Kehs. Beyond the normal maintenance of our current operating system, compiler, and graphics package, we are further developing the facility's work station and stereo viewing capability, and are developing a graphics package which is independent of any particular display terminal or interaction device. This will help us meet our goal of making our application programs usable on equipment other than our own.

Because our resource has been in existence for just 8 months, we are not yet prepared to undertake service or training activities. We would expect to begin developing these activities on a limited scale during our second year. In addition to this mission-oriented organization, there is a Resource Advisory Committee. Its members are Drs. Brooks, Eliel (UNC Chemistry Department), Hermans, Kim, Perl (UNC Physiology Department) and Richardson (Duke Biochemistry Department). Minutes of their meeting are attached.

II.A.2. Administrative Organization.

We are administratively part of the UNC Department of Computer Science, which is the home of most of the investigators and research associates and assistants, as well as of the graphics facility. As the resource develops and comes to the point of having systems ready for use by the community of users, administration of the resource will shift to the UNC Computation Center. This planned arrangement was discussed in our original proposal. At the moment, our current arrangement is quite satisfactory.

The Department of Computer Science's Computer Graphics Laboratory, the facility we use, is directed as a service facility by Mr. James Robb, the department's associate chairman. Under his supervision are Mr. Peter Nichols, an electrical engineer in charge of the equipment, and Mr. Wayne Babich, who provides user training and services.

Technical direction of the laboratory comes from the principal investigators using the Laboratory: Drs. Brooks, Foley, and Wallace.

II.A.3. The Resource Facility.

Figure 1 depicts the entire complement of equipment available in our resource, including the soon to be delivered Tektronix 4014 at Duke. The display console and most of the interaction devices are clustered around a work station, as shown in Figure 2. The work station design represents our first attempt at providing a convenient working area. Experience has shown that the keyboard and other devices are not positioned just right, so we are currently working on the design of a second station.

The various interaction devices in the drawing and photograph have been assembled to give us the widest latitude in designing the man-machine interface to suit the user. The various knobs and joysticks are useful for indicating positions, angles, sizes, or rates of rotation or translation. The manipulator arm, furnished by the AEC, will at some time in the future be used for positioning molecules or portions thereof.

II.A.4. Plans and Objectives.

We discuss in this section our short-term plans, and our long-term objectives. By the end of March, GRIP will be ready for extensive testing by Drs. Hermans and Ferro. They expect at that time to be working on two problems:

- Energy minimization of the protein rubredoxin and comparison of the results with the raw and refined conformations obtained by the crystallographers.

- A study of the complex formed by the active site of a serine protease (perhaps chymotrypsin) and the transition state of a substrate. These are advanced studies of complex structures and should provide many excellent situations in which to test the usefulness of the system. Such use will indicate what directions we should take in further enhancing the system's capabilities and usability.

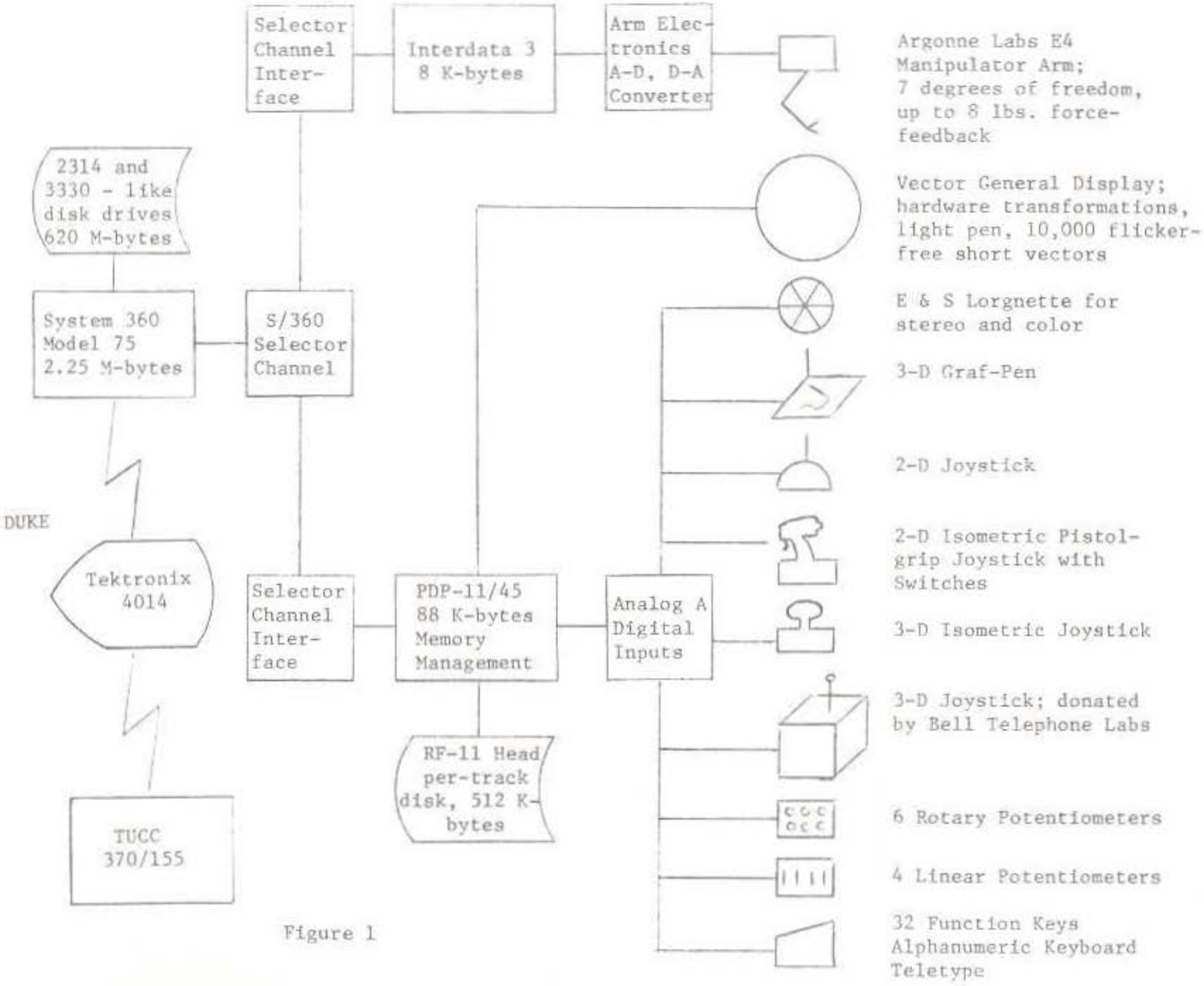
Selector Channel Interface
Interdata 3 8 K-bytes
Arm Electronics A-D, D-A Converter
Argonne Labs E4 Manipulator Arm; 7 degrees of freedom, up to 8 lbs. force-feedback
2314 and 3330 - like disk drives 620 M-bytes
Vector General Display; hardware transformations, light pen, 10,000 flicker-free short vectors
System 360 Model 75 2.25 M-bytes
S/360 Selector Channel
E & S Lorgnette for stereo and color
3-D Graf-Pen
2-D Joystick
DUKE
Tektronix 4014
2-D Isometric Pistol-grip Joystick with Switches
Selector Channel Interface
PDP-11/45 88 K-bytes Memory Management
Analog A Digital Inputs
3-D Isometric Joystick
3-D Joystick; donated by Bell Telephone Labs
RF-11 Head per-track disk, 512 K-bytes
6 Rotary Potentiometers
TUCC 370/155
4 Linear Potentiometers
32 Function Keys Alphanumeric Keyboard Teletype

Figure 1

Figure 2. Graphics Work Station

Any needed modifications will likely be completely ready for full use by summer. At that point, we plan to begin seeking out other potential users of the system.

Preliminary testing of GRAB by Dr. Kim will probably come in April. He will attempt to determine the structure of a model nucleic acid starting with the residue sequence and density map. He will likely use a combination of moving/rotating residues and bond twistings, as well as applications of REFINE. This will be the first real use of GRAB, and we expect that some aspects of the man-machine interface will not be satisfactory. We therefore plan for an appropriately modified second system to be ready during the summer. At that point Dr. Kim plans extensive use of the system. He will attempt to improve the structure of yeast tRNA PHE, starting with the current model based on 3-angstrom data [2]. We have also had preliminary inquiries from Dr. Lipscomb at Harvard. He would like to use the system this summer to study the enzyme aspartate transcarbamylase. Use by two active researchers will provide major guidance in setting the course for future work on GRAB.

GRIP and GRAB are both prototype systems. GRIP was built for the 11/45-Vector General using most of the code from Wright's earlier IBM 360/50 IBM 2250 system; GRAB is superimposed on GRIP. As a result, the organization of much of GRIP, and hence of GRAB, is based on design decisions made 5 years ago, for a system configuration quite dissimilar to our own. Despite this, in the summer of 1974, when our resource was first funded, we concluded that we should build onto GRIP to quickly learn more about the needs of biochemists and crystallographers. Choosing the alternative, starting from scratch, would have meant being less advanced than we are, without any guarantee that a new system would in fact meet the users' needs.

This strategy has been quite productive for us, and in retrospect was the right one. The clear implication of the strategy, however, is that we will at some point need to say "We have gone as far as is reasonable with GRIP and GRAB; now is the time to use what we have learned in designing a new system." Several inherent deficiencies in the current systems are apparent; others will doubtless be seen as time goes by. It is likely, we believe, that by late Summer or early Fall we will be ready to design a new system. It will be organized to meet the objectives delineated in our original proposal:

- It will be usable from a variety of graphics display terminals.

- If a powerful satellite (mini) computer exists, it will process all user actions not requiring extensive computations.

- For less-powerful satellites, some or all of the application program function will be easily movable to the main host computer. If there is no satellite, all functions will be performed by the host.

- Human factors (usability) considerations will be of paramount importance.

While the new system is being built, we will continue to expose biochemists and crystallographers to GRIP and GRAB.

The REFINE program is an integral and important part of GRIP and GRAB. Our collaborative work in this area is aimed at enhancing its function, decreasing the space and time needed for execution, and making it available to other users. Improvements of REFINE will be integrated into GRIP and GRAB, and used as we continue to verify its usefulness as a molecular model building tool.

We will implement a scaled-down version of GRAB for use with the Tektronix display in Dr. Kim's lab. Instead of displaying density contours on the screen, they will be on transparent sheets at right angles to the screen. The map will be merged with the molecule on the screen using a half silvered mirror.

Our facility work will continue. A new work station should be ready this summer. The device-independent graphics system (DIGS) implementation will start soon, so that later versions of GRIP and GRAB can use it. Additional stereo viewing techniques will be integrated into GRIP and GRAB. Some of the techniques will also be used with the Tektronix system.

II.A.5. Implications for future NIH Support.

Our work will not be completed by June 30, 1976, when our current funding expires. We will need continuing support to develop production programs, to engage in collaborative work, and to provide service and training to the general user community. As the resource moves toward providing routine services to users, a full-time resource manager will need to be funded. Indeed, such a manager could be

productive now, so we have been watching for a qualified individual. Unfortunately, very few people have the requisite management capability coupled with knowledge of biochemistry, crystallography, and computer science.

The display system being used for development work is owned by the UNC Computer Science Department, with the UNC Highway Safety Research Institute having a small interest. Thus it must be used for many purposes, as reflected on the form "Summary of Computer Resource Usage". As resource utilization builds up, we will have to buy more terminals. An ideal arrangement would be to have powerful refresh terminals available in Duke's and UNC's Biochemistry departments.

SUMMARY OF COMPUTER RESOURCE USAGE

Investigators	Department Institution	Project Title	Grant Support Identification No.	Agency	Dedicated Computer Hours	Time Sharing Computer
Core Research and Development Projects						
1. Hermans/Wright/Pique	Computer Science/ Biochemistry - UNC	"Graphical Manipulation & Refinement of Protein Structure (GRIP)"	1-P07-RR00898-01	NIH	352	(See Note)
2. Brooks/Kim/Britton	Computer Science/ Biochemistry - UNC/Duke	"Electronic Richards Box (GRAB)"	1-P07-RR00898-01	NIH	192	(See Note)
3. Wallace/Lipscomb	Computer Science - UNC	"Techniques for 3-Dimensional Viewing of Molecules	1-P07-RR00898-01	NIH	280	(See Note)
4. Foley/Wallace/et al	Computer Science - UNC	System Software Development & Maintenance	1-P07-RR00898-01	NIH	339	(See Note)
Subtotal: 4 projects					1163	
Collaborative Projects						
None						
User Projects						
None						
Training Projects						
None						
Non-Health Related Users						
Subtotal: 9 projects					1320	
GRAND TOTAL: 13 projects					2483	

NOTE: The Department of Computer Science pays on behalf of the UNC Computation Center partial rental cost on one disk drive. The Department rents two disk packs for its exclusive use and rents from the Computation Center a 512 K region in a 2-megabyte core memory. All of the above projects use this disk drive and disk packs and practically all projects involving Graphics run in the region reserved for the Department of Computer Science.

Grant No. 1-P07-RR00898-01
Activitiy: BRB
Date: 2/27/75

RESOURCE EQUIPMENT LIST

Equipment Located in Main Resource Area

Description/ Identification	Manufacturer	Type	Model	Date Installed	Date Accepted	Purchase Price	Annual Rent	Source of Funds
Graphics Display System	Vector General	2DR		5/20/73	8/27/74	68,311		NSF
PDP-11/45 - IBM 360 Channel Interface	Digital Equip. Corp.	DX11-EA		4/30/73	5/10/73	17,535		NSF
3-Dimensional Rotation, Translation, Scaling, Intensity Modulation, Post-Transformation Translation	Vector General	3 P3, IM1	PXYI	3/26/74	3/29/74	14,296		NSF
Communication Terminal	U.C.C.	Datel		7/14/71		4,120		NSF
CRT Communication Terminal	C.C.I.	CC/30		1969		8,000		NSF
Graf Pen	Science Resource Assoc.	3D		1969		4,000		NSF
Programmable Desk Calculator	Hewlett-Packard	HP-65		1974		854		NSF
32K Memory Unit	Cambridge Memory, Inc.			1974		8,043		NSF
Disk Controller	Digital Equip. Corp.	RF11		1973		5,200		NSF
Fixed Head Disk	Digital Equip. Corp.	RS11		1973		9,360		NSF
Interdata Computer	Interdata Corp.	#3		1969		20,000		Various
Master/Slave Manipulator	Argonne Laboratories	E3		1972		40,000		AEC
Lorgnettes, Interface Devices, Analog/Digital Converters, Analog Input Devices, Solenoid Drivers, Delay Devices, Alarm and Status Indicator Systems, Power Supplies, etc.	Various			Various		15,000		Various

Grant No. 1-P07-RR00898-01
Activity: BRB
Date: 2/27/75

RESOURCE EQUIPMENT LIST

Equipment Located Outside the Main Resource Area

Location/ Description/Identification	Equipment Manufacturer	Type	Model	Date Installed	Date Accepted	Cost Purchase Price	Annual Rent	Source of Funds
Storage Tube Display	Tektronix	To be delivered by 6/30/75				13,000		NIH

LOCATION: Duke University, Durham, North Carolina

II.D. SUMMARY OF PUBLICATIONS

Hermans, J., and McQueen, J. E., "Computer Manipulation of (macro)Molecules with the Method of Local Change". Acta Cryst. A30, 730-739 (1974).

IV.A. CORE RESEARCH AND DEVELOPMENT PROJECTS.

II A 1

IV.A.1. Stereo Viewing.

The tasks performed by the chemist in both the GRIP and GRAB systems are fundamentally three-dimensional in nature. The GRIP user manipulates the bonds of a molecule in a three-dimensional manner and asks questions of the system about sizes and 3-D directions of various computations/ results. The GRAB user attempts to twist and turn a molecular fragment inside a 3-D electron density map. In both instances it is essential that the flat image on the screen of the VG be given the illusion of depth. The GRIP/GRAB system provides 8 distinct techniques for accomplishing this. They may be applied separately or in various combinations.

In the following list, those techniques marked with an asterisk have already been implemented in GRIP and GRAB; the remaining methods will be available this spring. The implementation work is being done for the M.S. in Computer Science by Mr. James Lipscomb, under the direction of Dr. Wallace. Mr. Lipscomb is not being supported by grant funds.

Some 3-D effects may be achieved without employing a true 3-D technique. Then too, many individuals are unable to perceive true 3-D because one of their eyes is weak. It is necessary for them to use a monocular 3-D technique, i.e., one which provides the illusion of depth to a person who effectively has only one eye open. The Vector General display hardware provides the following 2 monocular features:

* Intensity depth cueing: the more "distant" features of a picture can be made dimmer than the nearer parts, "fading into the distance", as it were.

* Z axis clipping plane: A "plane of invisibility" can be drawn through the picture under the user's control. By moving the plane away from the viewer, successively more distant features disappear in depth order. The process can then be reversed.

Other monocular viewing techniques available rely on the ability of the mind to perceive an object as three-dimensional despite its being seen by only one eye (from one viewpoint) if only it is moved continuously. They are:

* Joystick control: The user can smoothly turn the picture by manipulating a joystick.

* Rocking: The picture may be set to automatically rock back and forth about an axis.

* Continuous rotation: For viewing purposes the image may be spun about a fixed axis at a constant rate.

A variety of true binocular 3-D viewing modes are available to those who want them.

* Lorgnette stereo: Two Evans and Sutherland lorgnettes have been interfaced to the system. Each has a rapidly spinning disk which alternately blocks the vision to each of the user's eyes. With the Vector General operating in synchronization with this device, each of the user's eyes can be shown a slightly rotated picture and true 3-D is effected.

* Side by side stereo: Two images are presented, one along side the other. The user can cross his eyes or look through a stereo viewer to see 3-D. This technique is akin to the "stereo pair" drawings which appear in virtually all chemistry research journals.

* Half silvered mirror stereo: One image is viewed directly through a half silvered mirror, the other is seen by reflection and appears to overlap it. Crossed polaroid sheets on the mirror and in glasses worn by the viewers direct the proper image to each eye.

These techniques can be applied in various combinations. We have observed that an intensity depth cued, Z axis clipped, joystick controlled, lorgnette stereo picture is perceived better than a picture presented using only one of the stereo techniques.

— (Move)

IV.A.2. Display-Independent Graphics System (DIGS).

IIA6

Many existing interactive graphics application programs (including ours) are written using a graphics subroutine

package which in one way or another mirrors both the exact capabilities and peculiarities of the display terminal with which it is used. This has very detrimental implications on our project's goal of making our programs usable from a variety of display terminals, ranging from low-cost, low-performance systems like storage tube terminals up to the high-cost, high-performance Vector General display which is the current heart of our resource.

Such a goal is a necessary one, because it is not reasonable to assume that all biochemists who want to use our programs will have Vector General 3DI displays with the exact set of options we have. A display-independent graphics package addresses this problem. Its design hides the details of any particular real display terminal behind the mask of a virtual display terminal, each generic characteristic of which can be realized either in the hardware of the more powerful displays, or in the software support of other displays. The point is that this mechanism hides display terminal details from the application programmer.

Display-independent packages are not new: several already exist, such as OMNIGRAPH [3], GPGS [4], and GINO [5]. We have examined these, and others, in detail. Unfortunately, none of them meets the highly-demanding needs imposed by our applications. Therefore we are undertaking the design and implementation of DIGS.

DIGS has a number of salient features. A full 2-D or 3-D scale-rotate-translate and window to viewport transformation is applied at the time the display code is generated. Extents (boxes) can be applied to portions of a picture to speed up the clipping process. Portions of the displayed image can be named, using a full block-structured naming capability.

The basis for picture modifications is the named segment, which is a collection of picture primitives such as lines, text, and points. The segment can be deleted, added to, or dynamically transformed. A viewport is a collection of segments. It too can be deleted, added to (by adding segments), and dynamically transformed. Thus a 2-level picture modification capability is provided.

Input is also device-independent, being handled through the mechanism of virtual devices. These include entity indicators (typified by the lightpen), position indicators (typified by a joystick-cursor), text string inputs (such as from a keyboard), and value inputs (as from knobs and dials).

The structure of DIGS makes it usable by displays ranging from storage tubes to Vector Generals. Our plans call for three implementations (each implementation includes substantial portions of common code) of DIGS; one for the 11/45-Vector General, one for a Tektronix 4014 coupled via communications link to a 360 or 370, the third for a 4014 coupled to a 11/45 and it in turn to a 360 or 370. Our status is that a preliminary architecture for most of DIGS has been written. The internal organization is now being mapped out in parallel with further refinement of the architecture. Major portions of DIGS should be usable by the end of this calendar year.

IV.A.3. GRIP and GRAB System

Work is proceeding in parallel on two interactive computer graphics systems, GRIP and GRAB, for displaying and manipulating molecular models. Both of these are evolving from the embryonic system (also called GRIP) which was developed at UNC in 1970/71. The present GRIP system is intended for use by the group of biochemists at UNC led by Dr. Hermans who is concerned with computational refinement of partially understood structures. The GRAB system is to be used by Dr. Kim and his group at Duke to facilitate the construction of molecular models from crystallographic data. Although the two systems are distinct, they share many routines and data structures, and most new facilities developed in either system are soon made available in the other. Thus, we encourage a cross-fertilization of techniques between the two groups of users.

At the beginning of the current grant period (6/30/74), we had in operation the original GRIP system suitably modified to run on the PDP 11/45 -Vector General satellite graphics terminal. This system was capable of displaying a detailed "stick" picture of a molecule, of calculating its internal potential energy and forces, and of executing precise manipulations specified by keyboard-entered constants and selections of light-buttons. Since then we have added the following facilities to form the current GRIP system.

- Dynamic control of the molecule's orientation, using a joystick.
- Stereoscopic display of molecules using an Evans & Sutherland lorgnette. This facility makes use of the work described in section IV.A.1 of this report.

- Selective display of individual residues in a molecule, and control of the detail displayed. Because a picture showing all bonds is confusing for most interesting macromolecules, a facility has been added for displaying three levels of detail: a set of lines connecting the alpha carbon atoms, all bonds of the main-chain of a molecule, or all bonds. These levels of detail are illustrated by pictures of the rubredoxin molecule in Figures 3 through 6. The level of detail displayed can be controlled independently for each residue. This can be seen in Figure 3 by the appearance of the side-chain attached to one residue and in Figure 6 in which all residues have been suppressed except for two shown with full detail.

- Manual analog specification of a position in the model space. A joystick with three rectilinear degrees of freedom can now be used to specify a point in the model space more quickly and more easily than with the keyboard entry and light-button facilities of the original GRIP system. Such input is often useful for moving a molecule to an approximate position.

In addition to these accomplishments, work is in progress on the following facilities:

- On-line refinement of molecular structure, as further discussed in IV.A.5.

- Manual analog twisting of selected bonds. The user will be able to select up to six bonds of his molecule and associate each with a knob. He can then twist these bonds by rotating the knobs and can observe the effect immediately in the display.

- On-line extension of system functions. Many of the current functions of the GRIP system are defined in terms of more primitive functions of the system. Facilities are being added to the system to enable the user to interactively define and test new high-level functions.

GRAB is an interactive computer graphic system with which a biochemist can determine the conformation of macromolecules from electron density maps. It is our Electronic Richards Box. The system architecture and interaction language are based on our analysis of a crystallographer's technique and principles of human factors.

Figure 3. Line connecting alpha-carbon atoms of Rubredoxin and one side chain

Figure 4. Rubredoxin main chain

Figure 5. Rubredoxin --all atoms

Figure 6. Two residues of Rubredoxin

GRAB is an extension of GRIP; hence, it includes all the latter's facilities. Additional capabilities are:

- Portions of the electron density map and molecule can be selected on-line for display.
- Electron density maps are presented as contours representing surfaces of equal density, as illustrated by Figures 7 through 10. Contouring is done on-line at up to three user-specified density levels. One or more of the levels can be selected for display. The contours for each level are prepared on three sets of orthogonal planes, any combination of which can be selected for display.
- Undesired contours can be removed individually.
- A single residue can be fit into the density map, in real time, using two separate three degree of freedom input devices, one for positioning, the other for orientation.
- Once fit, coordinates of any atom in the residue are readily available for examination by the user.

All the GRIP features currently considered at production status have been integrated into GRAB. A preliminary work station (Figure 2) is being tested. At present the density map facilities described above are all in operation, though we plan a more sophisticated mechanism for selecting which portions of the map and molecule to display. Residue translation and rotation are available. Algorithms for bond twisting are being developed now. The system has been tested with proteins, though only minor changes appear necessary to adapt it to nucleic acids. A version with the minimal features needed for production work is expected in April, with a final version completed during the summer.

Basic hardware and software facility support, as well as funds for batch and interactive computing, are necessary to carry out this research.

IV.A.4. Immediate Plans for the Tektronix Scope at Duke University Biochemistry Department

Our immediate plans for the Tektronix 4014-1 Graphic Display Terminal are divided into three stages. In addi-

Figure 7. A residue with no density map

Figure 8. A residue with 1 orientation of density map

Figure 9. A residue with 2 orientations of density map

Figure 10. Rubredoxin and its density map

tion, we will use it to access systems in use at NIH and Brookhaven. Initially we plan to use the display as a real time plotter to view various protein and tRNA molecules. Our first specific application will be to look at the refined coordinates of tRNA and to monitor the process of refinement in detail. At present we are doing this via a Calcomp plotter at Duke using the molecule drawing program ORTEP (by C. Johnson). We plan to implement a modified version of this FORTRAN program to run under TSO at Triangle Universities Computation Center (TUCC) so as to be able to easily view any portion (or all) of the molecule from any orientation. This will be done within a month of the terminal's arrival.

In the second stage we plan to implement a modified version of Jan Hermans' molecular model building and refinement program in an interactive environment including display capability such that we will be able to both manipulate and refine a macromolecular structure in real time. At present we are using a batch version of this program at TUCC.

The third stage will be to implement a semi-electronic or hybrid Richards Box on the display. We plan to use the display in an interactive way to view and manipulate a model, combining it, via a half silvered mirror, with real electron density minimaps in the form of Kodalith transparencies mounted on small plexiglass sheets. At present we are doing this in part; i.e., the minimaps are combined with static Calcomp plots of portions of the molecule. Alternatively, we might use stereo maps on microfiche, with a stereo viewer. At some point, as DIGS becomes available, this effort will merge with the GRIP/GRAB system. Initially, Dr. Sussman will be doing the programming described here.

IA 2

IV.A.5. Integration of REFINE into GRIP.

As was originally planned, we have integrated the manipulative approach used in the program REFINE developed by Hermans et. al. into GRIP. More will be said about REFINE in section IV.B.1. Using the method of REFINE, we have developed the following manipulative approach, to be used in building a model of a large molecule to fit other criteria (particularly an electron density map). Portions of the structure, such as peptide groups, side chains or parts thereof, are brought into the model space as rigid bodies and 'docked' into approximate position (see IV.A.3. for

methods). The pieces themselves have standard geometry, but it is clear that the action of moving them about separately has altered the geometry of the entire structure inside the model space to an unacceptable form. Once two or more disconnected pieces of a molecule have been positioned in the model, REFINE can be used to adjust the model so all bond lengths and angles (in particular, those at the connections between the pieces) are adjusted toward their ideal values.

We are at the start of executing a series of tests of this method (and comparisons with alternate ways of model building) with a cyclic decapeptide and various approximate representations of the conformation to be built. Examples of the approximate representations are another stick model, a stick model with errors and/or with missing elements, and a calculated density map.

Following this, we will incorporate the energy calculation portion of REFINE into GRIP.

IV.A.6. New Near-Neighbor Calculations.

II A 4

The GRIP/GRAB system will include programs for calculating the internal potential energy of a molecule and the forces acting on its individual atoms. The result from these programs can be displayed on command, and are also needed by the REFINE program. In theory, these calculations require that for each atom the distance to every other atom be calculated; the summation of certain functions of these distances is the energy. Due to the rapid drop-off rate of atomic forces, however, atoms far apart in a molecule have little effect on one another, and for practical purposes can be excluded from consideration. In particular, some value d can be chosen such that for any atom, only other atoms within distance d need be considered. Katz and Levinthal successfully approached this problem with "cubes" surrounding atoms. Jon Bentley has been investigating algorithms to find those "near neighbors" without having to examine all of the atoms in the molecule.

An algorithm that has excellent asymptotic performance has been developed. Letting n be the number of atoms in the molecule, for a fixed d the algorithm uses on the order of n log n distance calculations. All other known algorithms require on the order of n^2. The asymptotic expression assumes a very large n; however, the largest number of atoms

in any molecule we have investigated has been several thousands, and the mean number is much less. Hence the current effort is to make the algorithm timewise efficient for relatively small values of n. The algorithm has already been sped up by an order of magnitude since initial implementation, and is now competitive with other commonly used techniques. Bentley is currently implementing new speedups which may yield yet another very significant efficiency increase.

IV.B. COLLABORATIVE PROJECTS.

IV.B.1. The REFINE program

REFINE is a program package developed by Dr. Hermans in collaboration with Drs. McQueen, Wei and Ferro under support by a grant from the National Science Foundation. It is a batch FORTRAN program which can be used with sets of experimental coordinates to:

- test the stereochemistry of the coordinates.
- build a model best fitting the coordinates.
- build a model of low packing (or nonbonded) energy.
- find a structure of lowest energy, considering not only the packing energy, but also the energy required for deformation of the structure from ideal geometry.

The version of this program which performs the first two of these options was put into use in 1973 by Jensen and Watenpaugh at the University of Washington and by Sussman and Kim at Duke. In January 1975, a second version (REFINE2) was made available to the groups already using the first release and also to Feldmann at N.I.H. and to Meyer and Koetzle at Texas A & M and Brookhaven. A third version has been completed, and is used by Ferro and Hermans at U.N.C. A publication describing results obtained with model building calculations has appeared [1].

IV.B.2. Exportation of REFINE2

We hope to have REFINE2 accessible on the Crysnet system from Brookhaven. The group at Brookhaven is now working on conversion of the program to CDC-Fortran, and in May Hermans will spend some time teaching Meyer and Hanson (from Johns Hopkins) to use the REFINE2 program at Brookhaven. Feldmann is currently at work integrating REFINE2 into his molecular graphics system at N.I.H. This will involve much reprogramming, on which he will, if needed, consult with Hermans. We are following his work closely, and Feldmann's progress and

ideas will aid our own attempts at using REFINE as an element of GRIP.

IV.B.3. Use of REFINE3

IB

REFINE3 is the most recent version of our refinement program. This version minimizes not only the nonbonded energy of the model, but models deformations of the structure (such as bond angle bending) and distributes these properly over the molecule.

The program is being used in a series of calculations. The first of these, on cyclo-triproline, showed that excellent geometry is obtained with energy minimization of a strained ring structure. The second, on antamanide, a cyclic sodium binding decapeptide, showed that energy minimization of a preliminary x-ray structure gives improvement towards the structure obtained with classical crystallographic refinement. The first study is complete, the second nearly so. In the third project, recently begun, we intend to study the same problem in the refinement of a much larger structure, that of the protein rubredoxin. (This work is being carried out in collaboration with the group of Jensen, which determined and refined this structure). As a fourth project, we plan to study the activated state of the enzyme-substrate complex of a serine protease.

V. REFERENCES

1. Hermans, J., and McQueen, J. E., "Computer Manipulation of (macro)Molecules with the Method of Local Change", Acta Cryst. A30, 730-739 (1974)

2. S. H. Kim, J. L. Sussmann, et.al., "The General Structure of Transfer RNA Molecules", Proceedings of the National Academy of Sciences, USA, Vol 71, pp. 4970-4974, (December 1974)

3. Robert F. Sproull, OMNIGRAPH: Simple Terminal-Independent Graphics Software. Xerox Palo Alto Research Center, Report CSL 734, December 1973.

4. L. Caruthers, D. Groot, J. Patberg, General Purpose Graphics System, Katholieke Universitiet, Nijmegen, Netherlands, 1972.

5. Sancha et.al., GINO-F, Computer Aided Design Center, Cambridge, England.

APPENDIX A

MINUTES

Scientific Advisory Committee for Molecular Graphics Research Project

3:00 P.M. December 17, 1974 First Meeting

Present: Committee: F. P. Brooks, Jr., chairman, E. L. Eliel (demonstration only), J. Hermans, S. H. Kim, E. R. Perl
Research Team: J. D. Foley, V. L. Wallace, W. V. Wright, E. G. Britton, J. S. Lipscomb, M. E. Pique

Agenda:

I. Demonstrations of current status:

1. GRIP Molecule Display and Manipulation System: Wright and Pique
2. GRAB System for Fitting Molecules to Electron Density Maps: Britton
3. Lorgnette stereo: Lipscomb

II. Discussions of plans:

1. Direct-view storage tube at Duke
2. Cooperation with Molnar's team at Washington University
3. Next developments

Comments and discussions, organized by agenda topic:

I. 1. GRIP

a. Perl observed that whereas long menus serve as effective prompts for casual or infrequent users, for experienced users they represent an interference to effective use, because the user already knows the functions by heart. Hermans confirmed that he was experiencing this difficulty with the chief GRIP manipulation menu.

b. Perl inquired what facilities were available for off-line picture copy. Brooks responded that we kept a Polaroid camera. Perl said this is handy but not accurate enough, and that a research resource needs such a facility, especially for the user who comes from elsewhere and wants accurate snapshots of his results at any of several points. Perl's team uses a cheap Hewlett-Packard plotter for off-line hard copy. Wright explained that we have

in operation snapshotting to the 360/75 disk, but we do not at present have a plotting program that works from Hermans's representation in which the data is stored on the disk.

c. Perl feels that present methods of changing the scale are quite awkward and pointed out the desirability for zoom. Wright and others remarked on the difficulty of continual zoom, because zoom requires redoing the clipping. This cannot be done in real time with our system. Brooks remarked on the alternative of using the linear pot to specify scale and a PFK to indicate that the linear pot has been changed and should be read and the clipping redone.

I. 2. GRAB

a. Kim remarked on the desirability of moving the viewing window back and forth along the molecule during the fitting process. He explained that the usual procedure is to fit one residue at a time, working from some well-defined feature of the density configuration until the ambiguities have so accrued that one doesn't know what to do. Then he starts at a different well-defined feature of the density configuration and works along from there. Usually the residue sequence is known or postulated before the fitting process starts, but one moves back and forth from one residue to another to satisfy the constraints of bonding between the residues, while at the same time fitting each residue satisfactorily to its own local density configuration.

Hermans proposed that with a system such as this a best procedure would be to fit the residues individually and then run the Refine program to cause them to come into proper linkage with each other.

Kim responded that the display system is most useful in the early stages when one is beginning the fitting process. At this time one tries large-scale or even global changes in the position of a particular residue, or even occasionally changes in a residue sequence. The display gives most assistance in trying these large-scale changes. During this stage, however, the individual approximate fits of the residues are not close enough to permit the Refine program to link things together correctly.

Perl remarked that with real problems, one encounters some lacks of fit which turn out to be due to errors in the data, either in the postulated residue sequence or in the density data. This, he said, is why the graphical interaction is such a powerful tool: it brings suspicious areas and data errors more quickly to the investigator's attention.

b. Perl remarked on the importance of studying the protocols of real users as opposed to test cases invented by the system developers. Real users quite frequently turn out not to do what the system developers had expected and to want to do what the system developers have not expected, because of misunderstandings both about the users and about the problems. He related an incident in his laboratory where test cases with small amounts of data did not require a function and several weeks of observation showed no use of it. The function was taken out of the system. The first time the system was run with realistic amounts of real data, the function was badly needed. He suggested that quite early Kim, or one of his colleagues, use Britton's system to try to fit a real unknown set of data while we watch the details of the process.

Wright agreed, pointing out that his observation of Hermans at work has shown exactly the component of surprise. Even this past month the team has been surprised that Hermans preferred to refer to residues in Rubridoxin by number rather than by pointing to them.

Brooks remarked that we have in the past videotaped such test sessions and that the repeated study of a single videotape has proved to be very enlightening. We also log the user interactions on the disk and can print these for analysis. This is useful in conjunction with the videotape, especially when the user talks about what he is doing and what he wishes he could do as he works.

c. Perl again remarked on the long menus, which appear to be inflated with respect to the function used at any one time.

Brooks suggested that the on-line macro facility, which is in the plan, is a key technical approach to menu deflation. Analysis of the protocols, to define which functions are used in clusters, is also important for this.

d. Kim observed that his team now had available several sets of density data for which fits were not known, and that one of these could serve as an adequate starting test case.

II. PLANS

1. Duke DVST

The purchase order for the Tektronics is still awaiting formalization of the UNC-Duke contract. The Duke contract people have approved the draft and returned it to the UNC contract people for final formulation, about

last week. Foley will follow up with J. E. Leonarz. Our goal is to get the purchase order out before the universities close on Friday afternoon.

b. The present delivery time on the Tektronix DVST is about 4 months; we can therefore hope to have one by April. The plan is to operate it initially with the Tektronix Fortran program package running at TUCC. Our goal is to connect it to the UNC Model 75 using our device-independent graphics package. This will not happen until at least much later in the year.

c. Kim indicated his willingness to travel to Chapel Hill to use the VC facility for testing and experimentation since it will be ready for use some time before the Duke DVST. Britton is to solve the problem of parking space for Kim.

II. 2. Molnar's Team

a. Brooks reported on a conversation with Charles Molnar at Washington University. Molnar's plans to investigate structural reformulations of a hypertensive drug to reduce side effects was not funded. Hence, the macro-module system for molecular display was not torn down as expected and is still in operation. Brooks had called to inquire if Dave Barry might be available for a one-year visit at UNC, in the event molecular display work was not currently proceeding at Washington University. Molnar suggested that even though Washington's molecular display work was proceeding, such a visit might interest Barry, if someone at UNC is actively engaged in Carbon 13 Nuclear Magnetic Resonance work using Lanthanide probes. Barry is especially interested in this work and in the protein folding work of Hermans and might find a visit scientifically exciting. Molnar gave Brooks permission to speak to Barry directly if we can put together a proposition. Hermans reported that Karl Koehler in the Medical School is working on C_{13} NMR. Hermans doesn't know how serious this interest is. In addition, he reported that the Chemistry Department has recently acquired a C_{13} NMR apparatus, but does not know who the investigator is. He suggests that we check with Eliel. Kim thinks that such work is also going on at Duke and suggests inquiring in the Chemistry Department. Foley will follow on the matter.

b. The second matter discussed between Molnar and Brooks was the proposal by Molnar that UNC be a test site for one of the new macro-module molecular display systems now under design and construction at Washington University. He thinks this would be desirable from their point of view because of the different user community and the different technical approach being taken at UNC. He will prepare more specific suggestions on such cooperation. Foley should inquire again in about sixty days. In any event, Brooks thinks that we want to establish a closer and more cordial

relationship than the once-a-year site visit that has been the practice.

II. 3. Next Developments

a. Wright presented present plans for the extension of GRIP. A copy of his outline is attached.

b. Kim indicated the desirability of having direct manipulation of dihedral angles by knobs. He reported that one often wants to "stretch out" a molecule a little bit to make it fit a density configuration and this is done by manipulating the dihedrals. Britton will follow up with Kim. Wright understands the desired function.

c. Perl wants the user to have the ability to occasionally select from the collection of all functions that have been built in the history of the system. He feels one really needs a dynamic ability to group these functions together into menus. Wright reported that his plan is to substitute user selection of menus from the present PFK-controlled choice among fixed menus.

d. In discussing the use of the Refine program, Kim feels that one wants an indication as to which distances, angles, and dihedrals are farthest from their canonical values after each iteration.

Hermans says one really wants the program to either scan the molecules while looking at the criteria, highlighting the ones farthest from canonical values, or else to scan down the list of canonical values and report on the residues or atoms that are farthest away.

e. Wright inquired as to how often one needs to build up a structure from a collection of residues. Kim responded that one usually knows the sequence. Occasionally the postulated sequence is wrong. Even more rarely, it is unknown in advance, but facilities for on-line molecule construction from a group of residues are used so rarely that they should have low priority. He then described briefly the method he uses in fitting a postulated residue sequence to a configuration. In passing, he noted that it is not feasible with real models to go to a scale smaller than two centimeters per angstrom. This is as small as one can manipulate, because the inter-atomic spaces of finer scales are too small to get the fingers through.

Perl remarked that the process is precisely what one goes through in working a jigsaw puzzle. One starts with small separate clusters around well-defined features and then figures out how to make the pieces hook together to satisfy all the constraints.

Wright suggested that from the point of view of the Graphic Display System, the residue by residue fitting processes as Kim sees it is not really adding residues to the structure but instead one by one turning on the display of residues in the structure whose display was initially suppressed.

Kim indicated that the most common errors in the postulated sequence are in connectivity of whole sections. "Sometimes one will think this proline is 5 to 10 when really it is 45 to 50."

f. Perl inquired as to how much work a user can lose in a system crash and indicated the extreme importance of either an explicit save at rerun points or an autosave facility. In his installation they use the same disk storage routine that is used for plotting.

Brooks inquired as to whether we are now journaling on disks since going under CHAT. Wright responded yes, but since CHAT allows no MOD disposition our journaling survives only until the next session, so one must run a separate batch program to move it out of that data set if we want to save any particular session.

g. Perl indicated the importance of having an extremely simple but quite complete user manual, especially for clients working alone. Brooks wholeheartedly agreed and emphasized that user manuals neglect such pragmatic details as where to find power switches, where keys are, what phone numbers to call for interactive service, etc.

Brooks and Foley indicated that for the coming year we do not expect clients to have to run alone. We expect to be able to furnish system people to observe and assist clients during work sessions, even in the late hours of the night.

h. A long discussion followed as to how the user should orient himself with respect to the parts of a molecule that are outside the viewing window. There are two reasons why one may not see something on the screen. It may be logically suppressed by the selective display facilities, or almost all the components of the object may be suppressed because of the clipping algorithm. The latter case occurred during a demonstration today.

Foley suggested that the present gnomon (the rotated coordinate system display) be displaced or supplemented by a small figure of about the same size. It would show the complete alpha carbon backbone

and a dynamically maintained viewing box in proper scale and orientation. It would be desirable to be able to move the box in both size, location, and orientation using Noll's box and the Fist, exactly as Britton proposes to move the molecule with respect to the density map, but with supplementary controls for changing the size of the viewing box. He also suggested that the backbone display be coded in some way to indicate which portions of the molecule were and were not logically suppressed.

Wallace remarked on the possibility of using arrows on the screen to show which of the six faces of the viewing box have in fact been used to clip elements. The arrows indicate that something is not being displayed and lies outside the viewing box in the indicated direction.

i. Perl concluded the meeting with an invitation for the group to come see his display system. Brooks indicated his inability to do so before going on leave, but suggested that Foley, Wallace and others might very well want to do so.

The meeting adjourned at 5:45.

FPBJr:sem

12/17/74

GRIP Plans for First Half of 1975

William V. Wright

1. Satisfy the users' needs.

   The prime objective driving the development of the GRIP system is to satisfy the needs of the users. Thus, the choice of functions added to the system and the order of their development will be determined largely by the users and the course of their research. However, I currently believe the specific objectives listed below are the most useful that can be accomplished in the next six months.

2. Functions currently under development.

   a) Stereo display using an Evans & Sutherland lorgnette.
   b) Display of electron density maps as sets of contours.
   c) Six-dimensional manipulation of:
      i) entire molecular structures.
      ii) selected substructures using for manual input:
         i) Noll's box and the "Fist".
         ii) The AEC telemanipulator.

3. Functions for which development has not been begun:

   a) On-line definition of verbs including
      i) improvements to the definition language.
      ii) a text-editing facility.
      iii) a new interpreter or translator.
      iv) dynamic, programmed control of the menu of verbs presented to the user.
   b) New energy routines which efficiently treat only interactions between neighboring atoms.
   c) On-line specification of the chemical sequence of a molecule. [The user requirement of this function has not been clearly established.]

WVW:sem