Third Annual Report
Interactive Graphics for Molecular Graphics System

TR77-05

May 1977

Frederick P. Brooks, Jr.

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 00898-03.

13 May 1977

Table of Contents

List of Figures

NATIONAL INSTITUTES OF HEALTH
DIVISION OF RESEARCH RESOURCES
BIOTECHNOLOGY RESOURCES BRANCH

SECTION I - RESOURCE IDENTIFICATION

Report Period:		Grant No.
From: July 1, 1976 To: June 30, 1977		2-P41-RR00898-03
		Date of Report Preparation
		May 1977

Name of Resource	Resource Address	Resource Telephone No.
Interactive Graphics for Molecular Graphics System	273 Phillips Hall, UNC Chapel Hill, N.C. 27514	(919) 933-1074

Principal Investigator	Title	Academic Department
Dr. F. P. Brooks, Jr.	Professor & Chairman	Computer Science

Grantee Institution	Type of Institution	Investigator's Telephone No.
University of North Carolina at Chapel Hill	State University	919/933-2148

Name of Institution's Biotechnology Resource Advisory Committee:

Scientific Advisory Committee

Membership of Biotechnology Resource Advisory Committee: (* 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
Frederick P. Brooks, Jr., Professor & Chairman	F P Brooks	May 5, 1976

Typed Name & Title of Grantee Institution Official	Signature	Date
George R. Holcomb, Dean Research Administration-UNC at Chapel Hill	George R. Holcomb	May 16, 1977

II. RESOURCE OPERATIONS AND PROGRESS

A. DESCRIPTION OF RESOURCE OPERATIONS AND PROGRESS

General

This year we have operated with "maintenance" funding of $51,000, in fact lower than in any previous year, and a skeleton crew. This forced a choice between continued development of GRIP and its operation as an experimental pilot service. We chose to continue pilot service as our chief activity. We have made some minor modifications and improvements to GRIP-75 in response to our clients' immediate needs, but we have done nothing on the further development of GRIP-II (except write proposals).

Objectives and Operating Policies

We are building a comprehensive and effective interactive computer resource for seeing, manipulating, and computationally modifying mathematical models of complex molecules. We believe that our present resource has been shown to be as complete and useful as any in existence; we are aware of many inadequacies and needs.

Fundamental to our approach are the following objectives:

-- The GRIP system is designed to help chemists get results to their problems, 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 computation processes. That is, it is an aid to, not a surrogate for, human thinking and manipulation. Hence a strong emphasis on human factors research and on human engineering of the system.

-- GRIP is designed to serve many users, not one or two, so it must include 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. Hence when cuts were necessary, we cut development in order to keep on getting user experience.

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

-- who has a scientifically interesting problem, as assessed by our Resource Advisory Committee,

-- 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.

We have extended this offer publicly at the Columbia Workshop on Molecular Graphics (June, 1976), at the South Atlantic Protein Crystallography Workshop (April, 1977), and by word-of-mouth at other occasions. So far, we have had as many users as we have been able to handle.

Our users are almost exclusively working on the structures of molecules of considerable biochemical interest: proteins and nucleic acids. We aim to advance health-oriented biochemical research by enhancing the power of individual researchers through better tools.

Administrative and Scientific Organization

Figure 1 shows the planned organization chart for the project as proposed in Supplement II to our 1976 renewal proposal. Figure 2 shows how we have in fact operated this year. W. V. Wright is not supported with grant funds. F.

P. Brooks and J. Hermans also received no support in 1976-77, because of the funds shortage. All three were, however, active in the project in the roles shown.

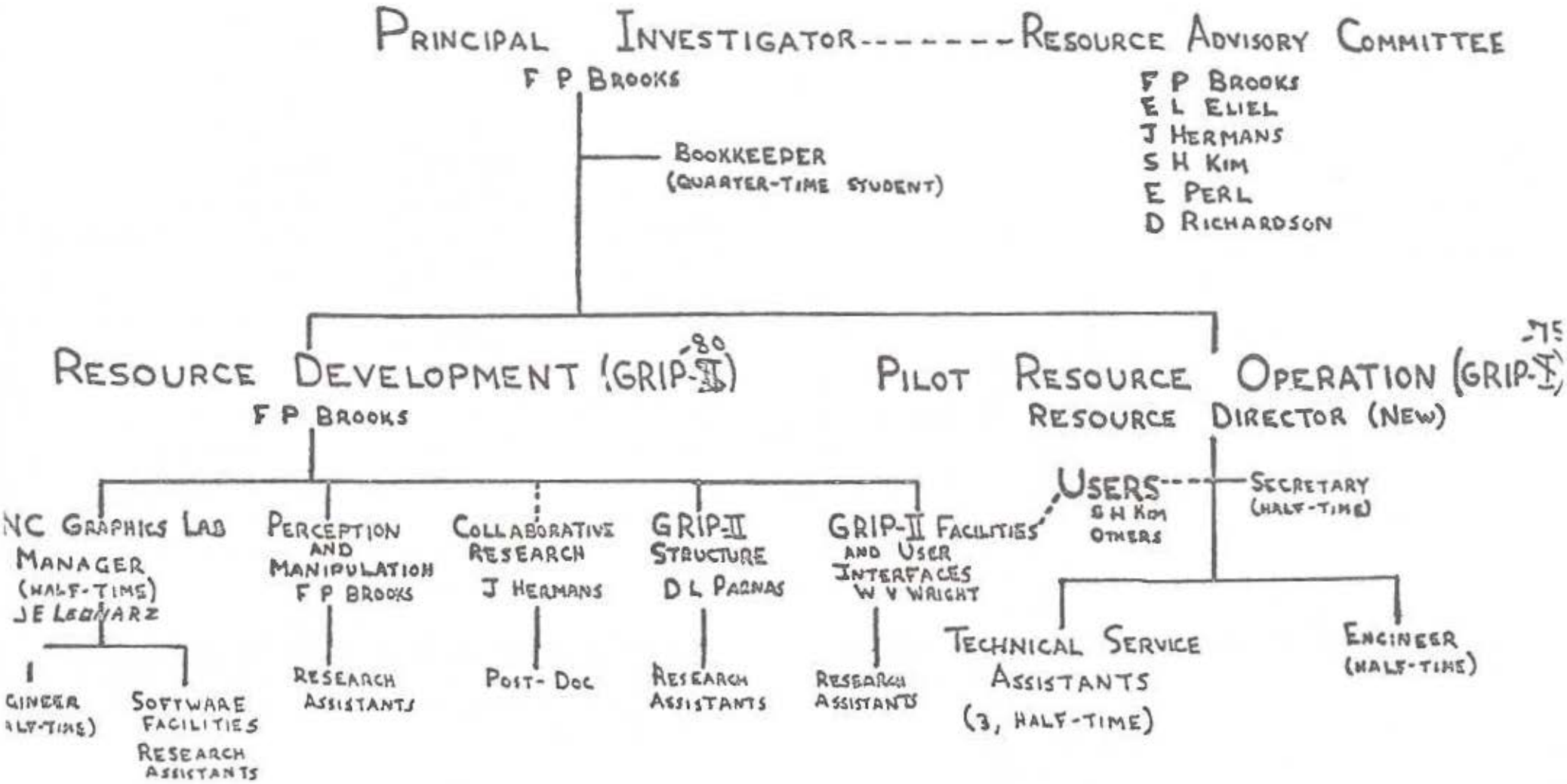

FIGURE 1. PLANNED ORGANIZATION OF UNC MOLECULAR GRAPHICS PROJECT

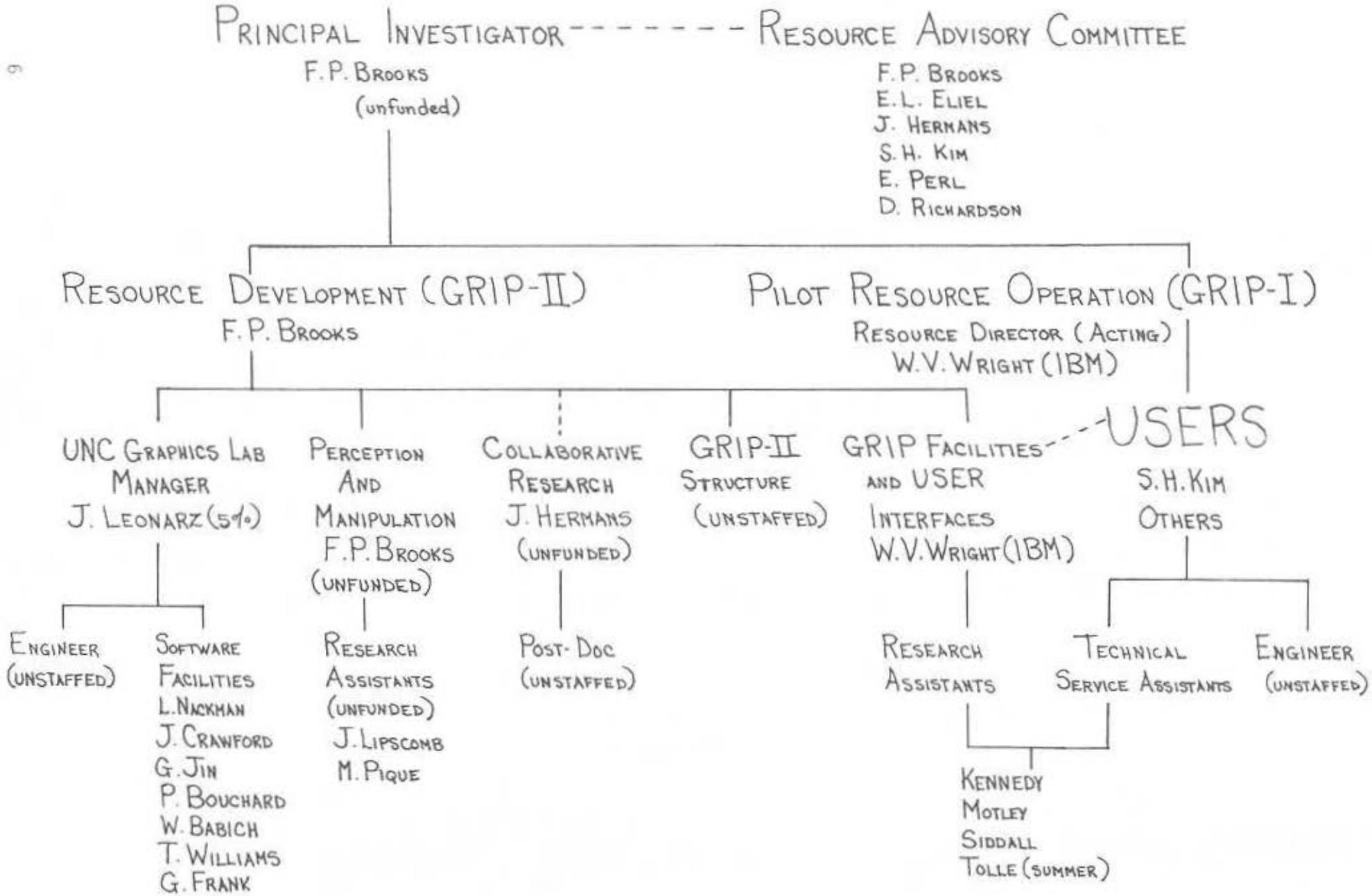

FIGURE 2. ORGANIZATION OF UNC MOLECULAR GRAPHICS PROJECT 1976-77

The principal changes in the administrative and scientific structure have been:

(1) The departure of Professor Foley, who was supervising graphics facility development and operation.

(2) The departure of Professor Wallace, who was supervising exploratory development of programming languages for interactive computer graphics.

(3) The graduation of Mr. E. G. Britton, who did a lot of the work on GRIP-75.

(4) The replacement of Mr. J. R. Robb as associate chairman by Mr. J. E. Leonarz.

(5) The coming to UNC of Professor D. L. Parnas, a renowned expert on software engineering. Dr. Parnas will join the project actively in the summer of 1977.

Core Research and Development

As mentioned, we did not progress on the development of GRIP-II, but we did develop a substantial number of enhancements to GRIP-75. There were generated in response to our users' most urgent needs as they came to light.

Most of these developments enhanced the facilities of managing the data libraries of many users, including facilities for bringing new data on line.

The other chief class of developments concerned molecule manipulation facilities, including on-line refinement and the building of initial models from molecular sequence and alpha-carbon coordinates.

Collaborative Research

As we served more users this year, we had more varied experiences, some quite successful, some unsuccessful. The list of users and a sketch of each of their scientific projects is given in Section IV.

The most dramatic success since our last report was accomplished by Prof. D. Tsernoglou of Wayne State University, who obtained in about 60 hours of system time a complete fit, ready for batch refinement, of a sea snake neurotoxin of 62 residues. This was almost an ab initio fit. They started with alpha-carbon coordinates read from a mini-map. No physical model or Richards box was built. This appears to be a first for a molecular graphic system.

The most serious failure was experienced by Dr. Delbaere and Mr. Brayer, colleagues of Prof. M. James of the University of Alberta. So we learned the most from their visit. We tried a new model-builder on them; it was clearly not ready for use, producing an initial configuration too far from the real one to be manually repaired in the time available.

The new technical issues raised by the Delbaere and Brayer experience have to do with when, how, and under what control departures from ideal geometry are to be allowed into the molecular model being fit to a density map. We have not finished our post-mortem analysis of the visit. Because the valence bonds of a real molecule are under stress and because of experimental error, the geometry of the molecular model which is the result of a fitting process is expected to depart slightly from ideal. Different users will want to inject these discrepancies at different stages in the fitting process and by different means. We learned from the Delbaere and Brayer experience that a widely-used system must accommodate the preferred methodology of all potential users. So a general system must offer choice in this matter, as well as in the others (perception, manipulation, contouring) we have seen earlier.

We append the letter reports written by our several visiting users. They indicate a great deal about the lessons to be learned from such collaborative research.

Service

The system has been available for molecular graphics use, both development and service, for 12 hours per day:

12-2 p.m.
5-8 p.m.
10-5 a.m.

This time has been substantially but not fully used, partly due to the inconvenience of some hours, but mostly due to our staff limitations on the amount of support we could furnish. The detailed data are given in Section II-B.

This year has driven home a major lesson about service. From our 1975-76 experience we believed, as stated in last year's report, that a key ingredient in our users' success was the steady availability of a computer scientist as "flight engineer," or "shepherd" to assist with all system problems.

Our 1975-76 experience showed a co-existence, if not a causative relationship, between shepherding and success. Our 1976-77 experience confirms our conjecture by showing a co-existence between poor shepherding and mediocre success. Because of the funding cuts, and the substantial number of visitors, we attempted to find lower-cost alternatives to full-shepherding of users.

First, our local users -- S. H. Kim and his group, J. S. Richardsor, J. Hermans, C. W. Carter, and J. E. McQueen -- had acquired enough experience to handle most sessions alone. The new library facilities permitted them to get by with even less help. This worked.

Second, with some of our visitors we tried the use of on-call, assigned shepherds. The users were to phone the on-call shepherd for advice in case of difficulty; for real trouble, the shepherd would come in. This didn't work; our users were too considerate, and accepted troubles and failure rather than phone a shepherd at 2:30 a.m.

Thirdly, several of our visitors were friends of local users, who volunteered to do the shepherding for them. This has had mixed results. The press of other duties has meant that some local users could not in fact stay up all night with their guests. Trouble was a result. In other cases, we think perhaps the shepherding user was somewhat partial toward the fitting techniques that had worked best for <u>his</u> problems, and did not encourage the visitor to explore the full range of technique choices available. In still other cases, the method has been very productive.

The lesson is not new but is very clear -- an effective service includes people as well as machines and programs. Our present opinion is that on-the-cheap shepherding is far too costly in facility waste and user time waste. Our new policy is to invite only so many visitors as we can give

full-service shepherding. For next year, that looks like one a month at most.

We are also developing much clearer guidelines about map quality, initial model data, user's previous experience, molecule complexity, etc. These enable us to give inquirers more precise and, we believe, more accurate appraisals as to whether their work is at a stage where our system will be useful.

Contrary to our expectations, our users have not been drawn chiefly from the nearby region geographically, but from the nearby parts of the discipline of crystallography. This has changed our concept as to what the client pool for our service should be, from regional to continental. It also forces us to emphasize new developments to minimize the stay in Chapel Hill of those who have come from afar.

This year we have had to choose between service and major development. We chose service, for the lessons to be learned. Now we must go ahead with development, and sacrifice service if something has to be cut. Else we cannot keep a first-rate development team together.

Training

Our chief training efforts have been in training users in system use and in training our own graduate students in the design and construction of man-machine systems. Neither of these efforts have been formalized; both have been very time-consuming.

Support

Most of the hardware used to implement GRIP-75 was owned by the Department at the beginning of the project. The salaries of our student research assistants, the charges for our use of the host computer, and maintenance of hardware were paid from a grant by the National Institutes of Health, #RR-00898. The services of Dr. W. V. Wright were made available under a joint study agreement with the International Business Machines Corporation, which was renewed in December, 1976, for the 1977 calendar year.

Plans and Implications for Future NIH Support

NIH's BRB originally stressed to us the importance of getting a usable, useful, and used system as quickly as possible and demonstrating chemically significant results. This has been done. The natural next steps are the construction of a product-quality, second-generation system and the establishment of a service-oriented multi-user facility. Each of these will take a lot of money. Both together will take upwards of $1 million over a three-year period. The real, tangible and usable results to biochemistry may well exeed, however, those from all of the millions heretofore spent on molecular graphics. The planned systems are not flashy or esoteric; they will, however, be carefully designed, user-oriented, and constructed to professional standards of quality. The probability of success is very high.

We have the experience, the team talents, and the user guidance to build such a major molecular graphics research resource. We are raring to go.

The plans are set forth in some detail in our renewal application of May, 1976, in the two supplements to it, and in our supplementary application of April, 1977. We will not repeat them here.

The next twelve months will determine whether or not this effort will be mounted. If it cannot be supported and mounted then, we shall have to turn our research efforts in another direction.

B. SUMMARY OF RESOURCE USAGE

Table I identifies the nine service projects which made use of the GRIP system during the past year. The first line of each project identification gives the last name of the principal investigator and the chemical substance(s) investigated. This project title is used to correlate the information in this table with the details of computer usage given in Table II and the narrative descriptions of these service projects in Section IV.

Tables II and III summarize the computer usage for all projects carried out under this grant. Computer usage by each service project for data preparation, data storage, and

interactive sessions is given in Table II. The resources required for all programming projects are given under system development in Table III. The batch computing time in these tables is reported as the number of minutes that were or would have been required on a S/370 model 165, the fastest of four computers available to us at the UNC Computation Center. Because our PDP 11 is always connected to the S/360 host throughout an interactive session, dedicated computer hours and time-sharing connect hours are the same. An analysis of 19 sessions showed that about three minutes of host CPU time are required for each session hour. This ratio was used to calculate the host CPU time for each service project. Host core memory utilization was also calculated from the total connect time for each project because 374K bytes of core are allocated to GRIP throughout each session.

We demonstrated the GRIP system to visitors to our department on about 30 occasions during the past year. These included:

Attendees and speakers of the Symposium on Structure and Dynamics of Macromolecules, Chemistry Department, UNC (March 17-18, 1977).

Dr. William R. Baker	NIH, BRB
Dr. Chris Bedell	Burroughs Wellcome Co.
Prof. Jonathan Hanson	Johns Hopkins University
Mr. Roland Ibbett	Manchester University
Prof. Thomas Isenhour	UNC, Chemistry
Prof. Daniel Jones, Jr.	UNC-Charlotte
Prof. Martin Karplus	Harvard University
Prof. Brian Matthews	University of Oregon
Dr. Martha Miller	University of Pennsylvania
Dr. Emanuel Piore	IBM Corporation

Table I: Identification of Service Projects

1. Carter -- HIPIP and Ferredoxin
   - Investigator: Charles W. Carter
   - Department: Biochemistry, Univ. of N. C.
   - Grant: (Unsupported)

2. James -- Alpha-lytic Protease
   - Investigator: Michael James
   - Department: Biochemistry, Univ. of Alberta
   - Grant: Medical Research Council of Canada 55-42289

3. Jensen -- Flavodoxin
   - Investigator: Lyle H. Jensen
   - Department: Biological Structure, School of Medicine, Univ. of Washington
   - Grant: NIH #AM 03288

4. Jensen -- Methemerythrin
   - Investigator: Lyle H. Jensen
   - Department: Biological Structure, School of Medicine, Univ. of Washington
   - Grant: NIH #AM 03288

5. Kim -- Phenylalanine tRNA
   - Investigator: Sung-Hou Kim
   - Department: Biochemistry, Duke University
   - Grants: NIH #CA-15802, NSF #GB-40814

6. Kim -- Alpha-lactalbumin
   - Investigator: Sung-Hou Kim
   - Department: Biochemistry, Duke University
   - Grant: (Unsupported)

7. Low -- Erabutoxin b
   - Investigator: Barbara W. Low
   - Department: Biochemistry, Columbia Univ.
   - Grants: NIH #NS-07747, NSF #BMS-73-01430

8. Richardson -- Superoxide Dismutase
   - Investigator: David C. Richardson
   - Department: Biochemistry, Duke University
   - Grant: NIH #GM-15000

9. Tsernoglou -- Sea-snake Neurotoxin
   - Investigator: Demetrius Tsernoglou
   - Department: Biochemistry, Wayne State University
   - Grant: NIH #HL-15958

TABLE II: COMPUTER USAGE FOR SERVICE PROJECTS

PROJECT IDENTIFIER	BATCH MINUTES ON S/370 MODEL 165	DEDICATED HOURS ON PDP 11/45 & CONNECT HRS TO S/360 MODEL 75 HOST CPU	HOST CPU MINUTES TO SERVICE INTERACTIVE COMMANDS	HOST CORE IN MEGABYTE -HOURS	HOST DISK SPACE IN TRACK -MONTHS OF IBM 3330 TRACKS OR EQUIVALENT
CARTER--HIPIP & FERREDOXIN	.54	15.5	46.5	5.8	---
JAMES--ALPHA-LYTIC PROTEASE	11.73	26	78	9.7	15
JENSEN--FLAVODOXIN	7.2	49	147	18.3	25
JENSEN--METHEMERYTHRIN	.37	112.5	337.5	42.1	40
KIM--PHENYLALANINE tRNA	.53	241.5	724.5	90.3	750
KIM--ALTHA-LACTALBUMIN	---	19.5	58.5	7.3	---
LOW--ERABUTOXIN b	15.9	42	126	15.7	15
RICHARDSON--SUPEROXIDE DISMUTASE	9.57	72.5	217.5	27.1	800
TSERNOGLOV--SEA-SNAKE NEUROTOXIN	9.76	96.5	289.5	36.1	40
TOTALS	55.6	675	2025	252.4	1685

TABLE III. COMPUTER USAGE SUMMARY

PROJECT CLASS	BATCH MINUTES ON S/370 MODEL 165	DEDICATED HOURS ON PDP 11/45 & CONNECT HRS TO S/360 MODEL 75 HOST CPU	HOST CPU MINUTES TO SERVICE INTERACTIVE COMMANDS	HOST CORE IN MEGABYTE -HOURS	HOST DISK SPACE IN TRACK -MONTHS OF IBM 3330 TRACKS OR EQUIVALENT
SYSTEM DEVELOPMENT	731.1	157	471	58.7	8765
SERVICE PROJECTS	55.6	675	2025	252.4	1685
DEMONSTRATIONS	---	42	126	15.7	---
TOTALS	786.7	874	2622	326.8	10450

Our research assistants spent a total of 4320 hours on all projects carried out under this grant. This time was used for preparation of the users' data, shepherding them through interactive sessions, and system development. Almost all of the system development carried out was done in reponse to the users' immediate needs and in thi sense was done in collaboration with them.

All service projects were health related, and we charged none of our clients for their use of the system. All local user groups (Duke and UNC) and Prof. Jensen, however, paid the UNC Computation Center directly for some batch processing of their data and its storage in our host computer. These direct payments to the Computation Center are not included in the financial data in the next section of this report.

Because we use our department computer graphic system for projects unrelated to this grant and because the system is sometimes down for maintenance, repair and system development, it is not always available to our clients. We have not determined the total number of hours available to our users, but the fact that they choose to use our system at inconvenient times such as the middle of the night indicates that at times our resources are nearly saturated.

C. EQUIPMENT

The resource continues to operate as part of the UNC Department of Computer Science, and it uses the Department's Graphic Facility. There have been no significant equipment changes in 1976-77; hence we do not repeat the configuration chart and list. This is a multi-purpose facility not equipped by NIH. It is used for several other research projects and learning activities in the Department.

D. PUBLICATIONS BY USERS

1. K. M. Beem, D. C. Richardson, and K. V. Rajagopalan, "Metal Site of Cu, Zc Superoxide Dismutase," Biochemistry, 16, 9 (May 3, 1977), pp. 1930-1936.

2. C. W. Carter, "X-ray Analysis of High-potential Iron-sulfur Proteins and Ferredoxin," Chapter 6 in Iron-

sulfur Proteins, ed. by W. Lovenberg (Academic Press, 1977).

3. S. H. Kim and J. L. Sussman, "Turn in a Conformational Pattern in RNA Loops and Bends," Nature, 260, 5552 (1976), p. 645.

4. D. C. Richardson, "Three-dimensional Structure of CU, Zn Superoxide Dismutse," in Superoxide and Superoxide Dismutase: Proc. EMBO Workshop, ed. by J. M. McCord and A. M. Michelson (Academic Press), in press.

5. J. S. Richardson, D. C. Richardson, K. A. Thomas, E. W. Silverton, and D. R. Davies, "Similarity of Three-Dimensonal Structure Between the Immuno-Globulin Domain and the Cu, Zc Superoxide Dismutse Subunit," Journal of Molecular Biology, 102 (1976), pp. 221-235.

6. J. L. Sussman and S. H. Kim, "Idealized Atomic Coordinates of Yeast Phenylalanine Transfer RNA," Biochemical and Biophysical Research Communications, 68, 89 (1976).

7. J. L. Sussman and S. H. Kim, "Three-Dimensional Structure of a Transfer RNA Common in Two Crystal Forms," Science, 192, 4242 (1976), pp. 853-858.

8. D. Tsernoglou and G. A. Petsko, "Three-Dimensional Structure of Neurotoxin a from Venom of the Philippines Sea Snake," Proc. National Academy of Science USA, 74, 3 (March 1977), pp.971-974.

9. D. Tsernoglou, G. A. Petsko, J. E. McQueen, and J. Hermans, "Molecular Graphics: Application to the Structure Determination of a Snake Venom Neurotoxin," submitted to Science.

10. D. Tsernoglou, G. A. Petsko, and A. T. Tu, "Protein Sequencing by Computer Graphics," Biochem. and Biophys. ACTA, in press.

Grant No. RR 00898
Section III-A

SUMMARY OF RESOURCE EXPENDITURES
BY BUDGET PERIODS

AF523

	Actual Previous	Current	Estim. Next
1. Personnel:			
a. Salaries & Wages	56,714.59	24719	
b. Fringe Benefits	3,011.82	325	
Subtotal	59,726.41	25044	
2. Consultant Services	-0-	-0-	
3. Equipment:			
a. Main Resource-Rented	-0-		
b. Main Resource-Purchased	-0-		
c. Supporting Equipment	-0-	4866	
d. Equipment Maintenance	2,481.80	7580	
Subtotal		12446	
4. Supplies	244.49	2455	
5. Travel	249.63	1000	
6. Alterations & Renovations	-0-	550	
7. Publication Costs	-0-	20	
8. Other			
a. Computer Services	2,937.68	9437	
b. Other	1,510.94	250	
Subtotal	67,150.95	51202	
9. Subtotal - Direct Costs			
10. Indirect Costs (Rate By Budget Period)			
a. Previous @ 48.13% S+W *	27,296.73		
b. Current @ 48.13% S+W *		11897	
c. Next @ unknown			
11. Total Costs	94,447.68	63099	

* Indirect costs based on period ending 6-30-72. Negotiations are continuing with DHEW over anticipated higher rate for more recent periods.

EXPENDITURE DETAILS
Direct Costs Only

	Current Budget Period		Estimate for Next Budget Period	
	% of Time or Effort	Amount	% of Time or Effort	Amount
1. PERSONNEL:				
1. Position Assoc. Chairman				
Names Leonarz, J.	5%	$919		
2. Position Graduate Assist-				
Names ants				
Bouchard, P.	18%	1920		
3. Nackman, L.	18%	1920		
etc. Jin, G.	18%	1920		
Crawford, J.	18%	1920		
Motley, R.	37%	3840		
Kennedy, G.	37%	3840		
Frank, G.	18%	1920		
Siddall, W.	30%	2920		
Babich, W.	13%	1000		
Williams, T.	13%	1000		
Summer Assts.		1600		
Subtotal - Direct Salaries		24719		
Fringe Benefits		325		
Total Personnel		25044		

EXPENDITURE DETAILS (continued)

		Current Budget Period	Estimate for Next Budget Period
2. CONSULTANT SERVICES		0	
3. PERMANENT EQUIPMENT:			
Main Resource - Rented		0	
1.			
2.			
3.			
etc.			
Subtotal			
Main Resource - Purchased		0	
1.			
2.			
3.			
etc.			
Subtotal			
Supporting Equipment			
1. Terminal	4866		
2.			
3.			
etc.			
Subtotal		4866	
Equipment Maintenance			
1. Contract	5847		
2. Other	1733		
Subtotal		7580	
Total Equipment		12446	
4. CONSUMABLE SUPPLIES			
(Group by Major Category)			
1. Electronics	1077		
2. Literature	434		
3. Photographic	60		
4. General	884		
5.			
6.			
etc.			
Total Consumable Supplies		2455	

EXPENDITURE DETAILS (continued)

	Current Budget Period	Next Budget Per
5. TRAVEL		
1. Brooks - England 236		
2. Brooks - Washington 64		
3. Brooks - Dallas 240		
4. Brooks, et al.- Rome,NY 460		
Total Travel 1000	$1000	
6. ALTERATIONS AND RENOVATIONS		
Total amount per period	550	
7. PUBLICATION COSTS		
1.		
2.		
Total Publication Costs	20	
8. COMPUTER SERVICES		
1. Type		
Rate		
Unit Used		
Total Computer Services	9437	
9. OTHER EXPENDITURES (Items not included in previous categories)		
1. (Telephone)	250	
2.		
Total Other Expenditures		
GRAND TOTAL - DIRECT COSTS	51202 (grant 51212)	

IV. CORE RESEARCH AND DEVELOPMENT

Serving Many Users

Much of the system development during 1976-77 was directed toward making it easier for many different user groups to use the system (one at a time but with interleaved sessions).

-- Facilities for accessing several libraries of electron density data, one for each group of users, were added to the system complementing the libraries of molecular models which were added the previous year. This required developing a new 4-fold-compressed format for the density data making it practical to store many sets of data on-line. The new format is also self-describing, which simplifies the handling of many different sets of data. Utility programs were built for translating density data from a variety of formats into our compressed form.

-- A new set of programs for preparing molecular models were developed so that our local users can manage their own data. These same programs simplify our staff's preparation of data for one-time users. They are implemented with an enqueuing mechanism which prevents conflicts caused by simultaneous access to any of the system data sets by the interactive system and one or more utility jobs.

-- We have developed mechanisms for avoiding most types of system crashes for which we know a solution. For unavoidable crashes, such as those caused by hardware failures and by other users of the host computer, we have developed a recovery mechanism which limits the loss for each crash to no more than a few minutes of the user's time. This mechanism also enables the user to recover from many of his own mistakes.

-- Several improvements were made to GTRAN, our utility program which translates definitions of the commands of our system from a special-purpose high-level geometric language into the form that is interpreted by the system during an interactive session. Use of this utility was further simplified by the addition of a catalogued OS/360 procedure for invoking it and by the creation of a standard dictionary of verb names. During 1976 we used this program to build new commands for invoking the sys-

tem functions, for constructing contour maps of density data, for manipulating molecular structures, for twisting a protein chain into an alpha helix, and for idealizing the geometry of these structures. Several of the newly defined commands were tailored to the requirements and methodology of individual users.

-- We have implemented a rudimentary password feature which can be used to block access by one user to another's data. So far none of our users has expressed a need to use it.

-- A facility has been developed to record the beginning time and duration of each interactive session. With a little cooperation from the users it also records who used the system and for what purpose. This facility can also be used to record dated, time-stamped comments. The system usage data given elsewhere were recorded in part by this facility.

-- Twice, in August and in March, the backlog of new facilities were integrated into the production version of our system. During the August integration we had to solve a substantial performance problem caused by the interaction of new facilities for stereo viewing and manual manipulation of a molecule. We also learned that we are very close to saturating the processing power of our satellite PDP-11/45.

New Application-related Functions

Several new functions were added to the system in response to the needs of our users.

-- It is now possible for the user with a single command to "connect" all of the variable geometric parameters of a residue to manual controls. This includes six parameters to position and orient the residue and an additional parameter for each twistable bond in the residue. A suitable manual device is used for each of these geometric parameters: a linear joystick with three orthogonal degrees of freedom controls the position of the residue, a three-axis joystick controls its orientation, and each twistable bond is associated with a knob that can be turned.

-- When idealizing the geometry of a molecular structure, the user can now specify groups of atoms that will be moved as rigid units. This speeds up the idealization because the optimization is carried out over fewer parameters. It also enables the user to selectively retain geometric relationships which may differ from the ideal geometry known to the system.

-- A new model-builder will now build molecular models from chemical sequence data and alpha-carbon coordinates. This model can serve as the startpoint for fitting operations.

Collaborative Projects

We consider all system development which stems from the use of our pilot system by chemists as collaborative work. Almost all system development during the past year was of this class. What we learned from our interaction with each of our clients is described in the following section on service projects.

Service Projects

1. Carter -- HIPIP and Ferredoxin

   Prof. Charles W. Carter used GRIP-75 to study the oxidized and reduced states of HIPIP and Ferredoxin. We also made a movie illustrating the activity of these molecules for him with the aid of our system.

2. James -- Alpha-lytic Protease

   In March 1977, Dr. Louis Delbaere and Mr. Gary Brayer working under Prof. Michael James used GRIP-75 to fit the first 46 residues of alpha-lytic protease to their 2.8Å map starting with a set of alpha-carbon coordinates taken from a mini-map. They not able to complete the interpretation of their map in the time available, and the partial fit obtained was not acceptable in quality. The starting configuration produced by our model builder was not good enough to justify continuing. We have improved our model builder since their experiment (see Low -- Erabutoxin b).

3. Jensen -- Flavodoxin

In December 1976, Prof. K. Watenpaugh, with the support of Prof. Lyle H. Jensen, used GRIP-75 to improve his interpretation of his 2.0Å map for Flavodoxin. Before using our system, Prof. Watenpaugh had obtained from his map coordinates for all non-hydrogen atoms of Flavodoxin without building a brass rod model of the molecule. Starting from this relatively advanced conformation he was able to obtain a satisfactory interpretation in about four days. Prof. Watenpaugh spent another four days on our system experimenting with its facilities.

4. Jensen -- Methemerythrin

In January 1977, Drs. Ronald Stenkamp (Molecular Biophysics, Yale) and John McQueen with the support of Lyle H. Jensen used GRIP-75 to fit a 113-residue subunit of methemerythrin to Dr. Stenkamp's 2.8Å map of this molecule. Approximately 70 percent of this structure is in four alpha helices. After unsuccessfully trying the method which had worked for Prof. Tsernoglou (see discussion below), this molecule was fit by first forming the four alpha helices with ideal geometry using an interactive command defined in our high-level geometric language. These were then manipulated into place using the manual fitting controls. Once this was done the "random" parts of the structure were successfully fit using the technique that had failed earlier.

5. Kim -- Phenylalanine tRNA

Drs. Joel Sussman, Stephen Holbrook, and Wade Warrant, and Mr. George Church working under the direction of Prof. Sung-Hou Kim continued the crystallographic refinement of the 3-dimensional structure of Yeast Phenylalaine Transfer RNA. A combination of molecular graphics and new refinement algorithm, the constrained-restrained least squares method, was used.

Prof. Kim's group also investigated three RNA-ligand interactions associated with this molecule:

(1) Metal binding sites in Yeast Phenylalanine Transfer RNA.
Drs. Sussman, Holbrook, and Warrant discovered why magnesium ions are essential to the activity of this tRNA molecule. By displaying a $F_o - F_c$ difference map on our molecular graphics system, they

were able to identify the number and the coordination geometry of the essential magnesium ions and study their specific stereo-chemical environments.

(2) Protamine -- double helix interaction.
Dr. Warrant investigated how Protamine becomes alpha-helical upon interaction with the tRNA and how it stabilizes the packaging of two adjacent double helical segments of tRNA's.

(3) Aromatic mutagen -- tRNA interaction.
Dr. Warrant identified tentative binding sites for several aromatic mutagens on the tRNA molecule.

6. Kim -- Alpha-lactalbumin

Mr. George Church has begun a crystallographic study of alpha-lactalbumin. He is attempting to fit the structure of lysozyme, a similar molecule, to his data on alpha-lactalbumin.

7. Low -- Erabutoxin b

In April 1977, Dr. Atsushi Sato (Chemistry, Tohoku University) and Mrs. Jane Richardson (Anatomy, Duke University) working under Dr. Barbara W. Low interpreted a 2.5Å map of the sea snake neurotoxin erabutoxin b. This is probably the same molecule that Drs. Tsernoglou and McQueen fit using our system in July and August 1976. Remarkably different methodologies were used by these two groups, however. Dr. Low's group started with a molecular model built from the residue sequence and approximate coordinates for the alpha-carbon atoms taken from a mini-map. Our model builder, improved since our experience with Prof. James's group, built a sequence of residues with ideal internal geometry and then positioned them in the model space using the given alpha-carbon coordinates. Instead of leaving these residues in the orientation in which they were generated, as our earlier program did, our present model builder makes use of the alpha-carbon coordinates for the adjacent residues to approximate the correct orientation for each residue. Starting with the resulting conformation Dr. Sato and Mrs. Richardson were able to manually fit the model to their electron density map. The resulting conformation will be used as the starting point for further refinement to this structure.

8. Richardson -- Superoxide Dismutase

Dr. David and Jane Richardson obtained a new map with 2Å resolution for their superoxide dismutase molecule and placed it in the GRIP system in June 1976. Fitting was continued from the conformation obtained with their previous map. An interpretation of one of the subunits of this molecule was finished, and the metal sites of all four subunits were compared.

Mrs. Richardson also used GRIP to make illustrations for one chapter of Principles of Biochemistry by White, Handler, Smith, Lehman, and Hill.

9. Tsernoglou -- Sea-snake Neurotoxin

In July and August 1976, Prof. D. Tsernoglou and Dr. J. McQueen fit a 2.2Å map of a sea-snake neurotoxin using GRIP-75. Their starting point was a set of approximate coordinates for the alpha-carbon atoms. The model builder of the Hermans/McQueen refinement program was used to create an initial conformation with approximately ideal molecular geometry. The principal method of fitting was to specify target positions in the density map for selected atoms and then to use the GRIP online routine for idealizing molecular geometry to move and twist individual residues until these target locations were achieved. The conformation thus obtained was sufficiently accurate to initiate a crystallographic refinement of the structure.

In November 1976, Dr. G. Petsko (Biochemistry, Wayne State University) used our system and an improved map to check the progress of the refinement and confirm the original fitting.

APPENDIX: CRITIQUES FROM GRIP USERS

WAYNE STATE UNIVERSITY

SCHOOL OF MEDICINE

GORDON H. SCOTT HALL
OF BASIC MEDICAL SCIENCES
540 EAST CANFIELD AVENUE
DETROIT, MICHIGAN 48201

DEPARTMENT OF BIOCHEMISTRY
313/577-1511

September 15, 1976

Prof. Frederick P. Brooks, Chairman
Dept. of Computer Science
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27514

Dear Prof. Brooks:

We want to express our gratitude to you and to the staff of the Department of Computer Science for making your computer graphics system available to us for the fitting of a molecular model to our electron density map of a sea-snake neurotoxin protein. Using your system to display the electron density map and a molecular model and then rotating the later until the best fit was obtained, one of us (D.T.), with the assistance of Dr. John McQueen of your University, was able to completely interpret the map in less than two weeks. We estimate that to do this by conventional (i.e., Richards' Box) techniques would have required two months or more. The atomic model obtained with the aid of your graphics system was sufficiently accurate to initiate a crystallographic refinement of our protein structure which has currently yielded an R-value of about 30%. To our knowledge this is the first time a protein structure has ever been interpreted entirely by computer graphics, without the building of a physical model at any stage. The ease and speed with which two inexperienced scientists were able to interpret the map bodes well for the future of computer graphics as a dynamic research tool. We were impressed by the care and thoroughness with which your system was constructed, but also by the need for continued input from protein crystallographers as regular users. An interactive system can only grow through regular dialog between designers and users, and we hope that you are given adequate support to refine this elegant tool, to produce an exportable version of it, and to make it available to some other users for testing and evaluation. We were sufficiently

convinced by our experience to conclude that henceforth all of our model building will be done by computer graphics.

Thank you again for making this opportunity available to us.

Sincerely,

Demetrius Tsernoglou

Demetrius Tsernoglou, Ph.D.
Associate Professor of
Biochemistry

Gregory A. Petsko

Gregory A. Petsko, D. Phil. (OXON)
Assistant Professor of
Biochemistry

UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195

School of Medicine
Department of Biological Structure

December 21, 1976

Frederick Brooks, Chairman
Department of Computer Science
University of North Carolina
Chapel Hill, NC 27514

Dear Dr. Brooks:

I would like to thank you and your colleagues at the University of North Carolina for a very productive and pleasant stay. Your generous offer of the use of the molecular graphics system was appreciated. The system was extremely useful in making the fitting of the protein, flavodoxin, to the electron density map both easier and faster than otherwise would have been possible. While it does have some limitations, they are primarily due to various minor hardware problems which in future improvements of the system can be readily eliminated. Conceptually, the design and implementation of the software appears to be well thought out and executed and will certainly serve as a model for future systems elsewhere. The adjustment of the molecular structure to the electron density map proceeded very smoothly from the rough initial set of parameters to what I feel is an improved new set. Since I was able to finish rather quickly, it allowed me several days to just experiment further with the system's limitations, what additional options might be useful and what I felt was the best ways for me to use it. I have discussed the results of my work with various persons involved with the system, but I would like to summarize some areas I feel improvements could be made.

1. More vectors should be able to be drawn in a given time span. This is especially true when using the stereo lorgnette.

2. Either more space should be made available on discs for larger electron density maps and/or finer grids; or the maps should be quickly calculated at selectable grid size in the region you are interested in studying.

3. A quicker method of removing undesirable contours that obscure the features one is currently studying could greatly increase its efficiency.

4. The current system requires an approximate starting structure.

It would be nice to be able to call amino acid or nucleic acid residues from a dictionary and build up a structure a priori.

My experience with your system has convinced me more than ever that computer graphics is an invaluable tool not only in the initial fitting of macromolecular structures, but also in their refinement and conformational study. Molecular graphics systems such as yours must be developed and supported to facilitate the research of macromolecular structure and function.

Thank you again for the opportunity to use your facilities.

Very sincerely yours,

Keith D. Watenpaugh

Keith D. Watenpaugh, Ph.D.
Research Assistant Professor

UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195

School of Medicine
Department of Biological Structure

Professor Frederick P. Brooks
213 New West
University of North Carolina
Chapel Hill, North Carolina 27514

January 5, 1977

Dear Professor Brooks:

This is in response to the copy of your letter to Dr. John McQueen concerning the fitting of a model to the hemerythrin electron density map by means of the display system you and your colleagues have developed. We will be pleased to cover the computing costs associated with this project. When these are known, please send the bill directly to me, indicating to whom or what project payment should be directed, and I will see that it is routed through the necessary channels here.

The hemerythrin structure is a large one, four subunits in the crystallographoc asymmetric unit with a total of 452 amino acids. The four subunits have been averaged by Dr. Ron Stenkamp, who is thoroughly familiar with the molecule and will be working with McQueen on the fitting. Thus, the actual fit will involve the averaged subunit with 113 amino acids. Because approximately 70% of the structure is in the form of α-helices, we are hopeful that a model can be fit to the averaged map in a relatively short length of time.

I want to express my appreciation to you in arranging for Dr. Keith Watenpaugh to fit flavodoxin on your display system, and now for Stenkamp and McQueen to fit hemerythrin. Having "hands on" experience is invaluable for our personnel, and we will try to transmit this, at least in part, to our colleagues at the American Crystallographic Association Meeting at Asilomar next month.

With best wishes,

Yours sincerely,

L. H. Jensen, Ph.D.
Professor

LJ/hg

cc WVW

(over)

Yale University *Box 1937 Yale Station, New Haven, Connecticut 06520*

DEPARTMENT OF MOLECULAR BIOPHYSICS
AND BIOCHEMISTRY
(203) 436-4817
Jan, 21, 1977

Dr. Fred Brooks
Dept. of Computer Science
Univ. of North Carolina
Chapel Hill, N.C. 27514

Dear Dr. Brooks:

I want to thank you for allowing me to work with Dr. McQueen fitting methemerythrin on your graphics terminal. We managed to fit the alpha helices, but had trouble fitting the random chain containing the first 20 residues. We were't sure where the polypeptide chain actually went, so after returning to New Haven, I fit the map with a labquip model and sent the atomic coordinates to McQueen so he can fit the model to the density using the graphics set-up.

I learned a lot during my week. Enclosed is a running account of our progress, problems, and what things I was impressed with.

Please express my thanks to Dr. Wright and the rest of your staff for their help, and especially doing the extra work necessary for us to twist up the alpha helices. Getting that instruction installed was the difference between getting some structural work done and just looking at your facility.

Sincerely,
Ronald Stenkamp
Ronald Stenkamp

Monday, Jan. 10, 1977

Went to the graphics lab and looked a bit at the map. The hemerythrin map looks very nice. Should be fun to fit it.

Tuesday, Jan. 11, 1977

Worked on the graphics system from 10 last night to 4:30 this morning. It's a lot of fun. Found some beta carbons for McQueen to use in making the model. Sure seems like you need a lot of information before using the terminal. I could have gotten the beta carbons off the plastic map in 30 minutes because I could move the map around faster by hand.

Here are some points from my first night.

1. Spends a lot of time contouring. Need to speed it up.
2. Took 6 hours to find some beta carbons. I was just learning, but seeing the whole map at once makes it a lot easier to do things.
3. You really want to see more of the map at once. You need the plastic map there to get a general view. But I thought the purpose of this was to get rid of plastic maps. Hope it does better for fitting.
4. Seeing the map contoured in three directions is nice.
5. The system needs a write-up for temporary users.
6. I used the twisting joy stick a lot. Wanted to twist the handle, though.
7. Options used and whether they're useful.
   a. CONON and CONOFF. OK for adjusting directions of contouring.
   b. Various things entering coordinates with the light pen. There must be an easier way to read in data.
   c. CEN-POS, CENXYZ. OK.
   d. POSXYZ. OK for getting coordinates out.
8. Used the 3-d positioner. It's OK. I lost track of the cursor in the map a couple times.
9. Brightness and depth queueing. Haven't caught on to them, yet.
10. Spinning and rocking weren't used except as toys last night. I'd much rather use the joy stick.
11. The system drastically needs a map cutting ability. You need to be able to contour just a few sections at a time. Something faster than REMVCON would be helpful. Maybe some sort of cutting plane or sphere could be put on the screen to cut out large regions. If you could remove areas of the map, you wouldn't have to remember what vectors are connected to make one contour.
12. I'm not at ease with the light pen. You end up moving your hands from device to device. I'd rather enter commands from a keyboard or buttons.
13. Using DELETE to clear the stack after REFINE is a pain. Takes too much time.

Worked from 5 to 8 on the graphics system. It's working OK. Went back to work at 12 midnight. Met Jim Lipscomb and Mike Pique. They were rather taken aback when I told them I hadn't done anything yet that I couldn't have done faster on the plastic sheets. I also complained about the light pen.

Wednesday, Jan. 12, 1977

We spent all our time yesterday and this morning just trying to get the carbonyls pointed the right way along the helices. Haven't finished that yet. Have done no fitting. There are two approaches to fitting maps. The one espoused by McQueen is to spot targets in the density and use their idealization program to pull the molecule into the right place with ideal geometry. The other is to cut the residues loose and move them around with another joy-stick-like thing to get them in the density. Then the thing can be idealized later. Neither method is too swift for this project because of the crazy way you have to get your molecule into the machine. Herman's model building procedure didn't work too well for this molecule because the alpha carbons were too rough. The model that's displayed is grossly nonideal. Some residues have Calpha, C' and N all

at the same location because the program allows enough distortion to do it that way. This means that the method of physically moving the residue can't work since the residue is sterically awful. McQueen's method suffers in that all those targets need to be located in the density. The set up time is quite large. According to Pique and Lipscomb you could start with an extended chain with absolutely ideal geometry and then fold it into the map. That might have been slightly better, but if I were doing it with a model, I'd build some alpha helix first and then fit that whole unit to the map. I would think that being able to use larger building blocks should have been made easier. Anyway, I hope I get to fit some density sometime.

I had trouble trying to see where the hydrogen bonds in the helices were. It's not all that easy to see what the stereochemistry is. Maybe I need to play more with the depth queueing, but what little I did try of that didn't help much. I'm also having problems with inverting the structure. Can't tell the front from the back. Also, if I use the depth queueing, I sometimes want to put the brighter lines at the back instead of the front.

Had more trouble with the light pen. Not steady enough. It moves too easily from item to item on the screen.

I'm becoming quite pessimistic about all this. Seems very expensive for what I've gotten from it so far. It's much easier to put your hands on a model and move it around. I really need to use the residue fitting option and see if that's what I like.

New developements. Wright has gotten McQueen's old REFINE up so now that'll go faster. We're starting over with an extended ideal chain that we can fold up. Having to add an old command to the machine. If this all works out, we should be getting some stuff done now.

Thursday, Jan. 13, 1977

Got more done last night than we had in the other two nights. Started over with an ideal extended chain, folded up the helices using the four year old instruction that had been discarded, moved the helices into the density and did some small rotations and translations to fit the density better. We didn't use the flying adjustments too much. They're a bit hard to zero out so they stay still. We just picked axes and did small rotations. After the helices were set, the other pieces were moved to close to where they should be and were then moved individually to the density. After this, the side chains were rotated to fit the density better. Haven't finished that yet. Normally, the pieces would be fine tuned using REFINE, but we're running out of time. Today, we'll work on the non-helical parts so John can continue after I leave.

I was impressed with how easy it was to change the instructions on the scope. Also, it's impressive how the operations you have to invoke to do things make sense if you've thought about moving models around. The computerese hasn't hidden the basic things you need to do.

It's hard to describe all the operations we found useful. Here's a try.

1. contouring routines
2. turning off sections along each axis
3. removing contours to clean up the map
4. scaling of the picture
5. windowing to fill the screen
6. showing alpha carbons, backbone or all atoms in the model
7. centering on position of cursor, residue, xyz, or atoms.
8. defining axes and roatating about them
9. moving pieces of the molecule
10. freeing up the torsional angles of each residue

We haven't used stereo yet since John can't use it. Maybe I'll ask to see it in operation tonight.

I'm not sold on computer graphics yet. There are enough little things bothering me, like flicker, small number of vectors, inverting the map and using the light pen, that I'm not sure I wouldn't rather use a model to fit things. Also, the economics of it bother me. A good terminal must cost $200,000. If in a Richards box you can fit 1 residue an hour, for 480 residues it takes 480 hours or 120 day (since you can only work so long per day). At $10,000 a year for salary and $10 per brass residue, that's $3000 + $4800 = $7800 for a model containing 480 residues. McQueen said you could only fit 1 residue per hour on the terminal, but with a fit comparable to the Richards box, without doing the idealization, you could maybe get 3 or 4 residues per hour. Then 480 residues take 30 days at 4 hours per day. So it costs $1000 in salary just to run the terminal. If you use your old model parts, the difference in price between computer graphics and Richards box is $2000 for 480 residues. You can build a lot of model for $200,000. So the reasons for computer graphics are convenience (for some) and the increased flexibility (You can use it for other things besides fitting models if you want) You know, this is really incredible. There's no lab in the country that'll solve 100 structures. For most labs, it would take 200 years to solve that many structures. I better check my estimates of rates of fitting. In any case, there's room for a couple factors of two in this calculation before the machine is worth it.

I must not be modern enough. Part of the desire to get a graphics system is the thrill of being up-to-date and at the forefront of technology. I can't buy that. The other argument is the increase in speed. I'm not that impressed with factors of 3 in time when it costs so much to save it. I guess I'm not sold on high technology just for itself. It's good that it's being developed, but not everyone should buy one if it's not practical.

While waiting for John this morning for breakfast, Dr. Brooks came along. I asked him about the economics. He says $150,000 is a better estimate of cost of equipment. Also, Fetsko and Sung Ho Kim claim fitting rates of 2 to 10 per hour. He also mentioned that they want to develop another system and set up a resource center for doing graphics. Here we go again with high energy physics-like facilities. Ugh, those sorts of facilities are not the neatest way to do science.

Friday, Jan. 14, 1977

Didn't work too long last night. Got tired about 3:30. Couldn't fit the first 20 residues. Not easy to see where the chain goes on the terminal because you can't see the whole thing at once. I'm going to plot that part up at a bigger scale, fit it with labquip stuff and send John the coordinates. Then he can fit it. I did get to use the stereo option. The lorgnette is not too neat. The stereo isn't that good and looking through the little plastic thing is a drag. It also limits how much you can see because of the reduced number of vectors in each view. The side by side stereo had a lot of flicker. I see why Sussman said he could work only for a few hours at a time before getting a headache.

DEPARTMENT OF BIOCHEMISTRY

THE UNIVERSITY OF ALBERTA
EDMONTON, CANADA T6G 2H7

April 4, 1977

Dr. W. Wright
Department of Computer Science
University of North Carolina
New West Hall
CHAPEL HILL, North Carolina 27514
U. S. A.

Dear Bill,

I want to thank you very much for the hospitality that you extended to Louis and Gary during their stay at UNC in your department. Not only did they enjoy visiting your group but also they profited immensely from the "hands on" experience with the graphics system.

They have written up a brief report on their visit with some critical comments from a user's viewpoint which I hope will help you in your continuing development of the system. I certainly was not cognizant of the problems involved in the graphics field and the algorithms must be horrendous. I would like to give a "long distance" impression of what a crystallographer would like to have, given three things to start from a) an amino acid sequence, b) an electron density map and c) a set of rough starting coordinates from a small scale map which is interpretable in terms of α-carbon coordinates (and perhaps selected side chain atom coordinates that are recognizable).

The basic philosophy regarding electron density map fitting that I would follow, perhaps because I am most familiar with that thesis is a combination of Richards' Box with the Diamond model building program (Class II). Firstly to have a dictionary (as you do) of standard geometry side chains and peptide links. Building could of course start at either end of the chain but to be consistent perhaps to use the strategy of Diamond and build (or fit) from the C terminal. Therefore I would initially fit the terminal α-carbon and carboxyl group with the side chain having at my disposal as variables ψ, and side chain χ angles. Then to add on (after establishing a reasonable fit of this terminal group) the penultimate residue (n-1) with its peptide link, side chain, α-carbon. At this stage I would use the chemical data that dictates residues n-1 and n are joined by a planar trans peptide bond and I would preserve that geometry until strain dictated that something had to be altered (eg. τ at the α-carbon).

The addition of the penultimate peptide and residue could be positioned anywhere with the only requirement that the nitrogen be bonded to the α-carbon of residue n and have the proper geometry (distances and bond angles).

Page 2. ...

Then, presumably (hopefully) one could achieve a reasonable fit of peptide bond, α-carbon and side chain with simultaneous (or sequential) adjustment of ϕ_n, ψ_{n-1} and χ_{n-1} dihedral angles. This then would constitute a cyclic procedure until one arrives at the N-terminus of the polypeptide chain.

Clearly strain would have to be allowed for in order to have the best fit possible. Strain could be introduced at the level of the τ (tau) angles at α-carbon atoms in order to maximize model/map overlap. I would not vary τ unless it was a last resort however and from our experience these angles differ by <15° from their accepted 109.5°.

The above paragraphs you may conclude are ramblings of a crystallographer. However I hope that you may find it helpful as well as the included comments of Louis & Gary. They were impressed by your system and would have liked to stay there to complete the job. I enjoyed the polaroid stereo views very much. We shall complete the α-lytic map with a Richards box however as I cannot afford another trip, nor can you afford to have us there I am sure. Thank you very much for your hospitality. Please give my very best regards to Professor Brooks.

Yours sincerely,

Mike James.

Mike James,
Associate Professor

MJ/dh
Enclo.
c.c. Prof. F.P. Brooks

COMMENTS ON *GRIP*

The features of the UNC GRIP system which we liked were the following: the toothpick for rotating the map and the structure together gives a very good three dimensional effect and is excellent for fitting individual side chains; the intensity variation with depth also helped give a three dimensional impression; the graphics displays of the model after one has built the structure are excellent; one can rotate the molecule into many different orientations and display certain residues on the whole backbone.

The features which we didn't like were the following: the initial peptide linkages were essentially random as to distance, angle and planarity; it was difficult to detect the carbonyl oxygen "bumps" in the electron density map probably due to contour level height; one must contour the map section of interest each time on line; one should be able to rotate structures in all orientations - there was a cone of space which was inaccessible due to the physical limitations of the toothpick.

Our suggestions are to start with ideal geometry as much as possible from a model building program. The positions of the C-alpha atoms and possibly the C-beta atoms for each residue could be measured initially from a mini map. These coordinates could be used in a "Diamond" -type model building program; the τ (angle N-C_α-C') values may be allowed to vary in order to least-squares fit the model to the original measured coordinates. The idealized geometry of the peptide bond should be maintained throughout the map fitting as much as possible. Fitting of the model to the map could proceed as does Diamond's model building program i.e. from the carboxyl terminus rotating around ψ to fit the side chain and the N-atom and then fit the planar peptide unit by rotating about ϕ. The next residue would be fit by fixing the position of the previously placed peptide bond and rotating around ψ and then ϕ in turn. Refinement such as that which is currently available in your system could be carried out in a molten zone of $\sim$3 residues.

It would also be convenient if the electron density map was contoured in three different levels on disk-one could call in the map section desired and work at the middle contour level, proceed to the higher level if desired or to the lower contour level for fitting carbonyls and detecting hydrogen bonding.

There should also be an option to use non-orthogonal maps. Also labels for ambiguous atoms of side chains such as O in threonine, N in histidine and N, O in asparagine and glutamine would be helpful.

DEPARTMENT OF BIOCHEMISTRY

THE UNIVERSITY OF ALBERTA
EDMONTON, CANADA T6G 2H7

April 5, 1977

Dr. W.V. Wright
Department of Computer Science
The University of North Carolina
at Chapel Hill
New West Hall 035A
CHAPEL HILL, N.C. 27514
U.S.A.

Dear Bill:

Thank you very much for your hospitality and your assistance when Gary and I were working with you on your GRIP system at Chapel Hill. We found the experience to be very informative and certainly most worthwhile. We have compiled a list of our comments on your system and they are being sent to you under separate cover by Dr. James; we hope that these comments will be of help to you in the further development of your system.

We also hope that sometime in the future we will be able to collaborate with you again in this manner.

Please give our regards to Professor Brooks.

Thank you again.

Yours sincerely,

Louis Delbaere

L.T.J. Delbaere

LTJD/dh

College of Physicians & Surgeons of Columbia University | *New York, N.Y. 10032*

DEPARTMENT OF BIOCHEMISTRY

630 West 168th Street

April 27, 1977

Professor Frederick P. Brooks, Jr.
Department of Computer Science
University of North Carolina
New West Hall
Chapel Hill, North Carolina 27514

Dear Professor Brooks:

Dr. Sato has returned here from working with Mrs. Richardson at Chapel Hill. I am happy to take this opportunity to thank you for your helpfulness. It was a boon that we were able to use the computer graphics system to interpret our map during the short time Dr. Sato is here in this country.

Jane Richardson certainly knows of the proper form of acknowledgement when we write up the results of these studies. I am most appreciative. Please do let me know if there should be an occasion on which I or my laboratory might be helpful to you or one of your colleagues.

Yours sincerely,

Barbara W. Low.

BWL:rd

Barbara W. Low
Professor of Biochemistry