Second Annual Report
Interactive Graphics for
Molecular Graphics System

TR76-04

April 1976

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

SECTION I - RESOURCE IDENTIFICATION

Report Period:	Grant No.
From: June 30, 1975 To: June 29, 1976	5-P07-RR00898-02

Date of Report Preparation

April 1976

Name of Resource	Resource Address	Resource Telephone No.
Interactive Graphics for Molecular Graphics System	273 Phillips Hall. UNC Chapel Hill, N.C. 27514	-
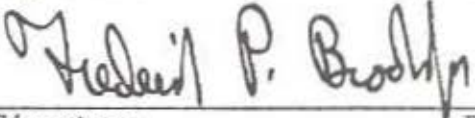
Principal Investigator	Title	Academic Department
Dr. F. P. Brooks, Jr.	Professor & Chairman	Computer Science
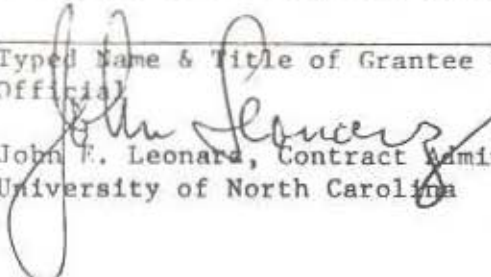
Grantee Institution	Type of Institution	Investigator's Telephone No.
University of North Carolina at Chapel Hill	State University	919/933-2148

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		May 5, 1976

Typed Name & Title of Grantee Institution Official	Signature	Date
John F. Leonard, Contract Administration University of North Carolina		

1

TABLE OF CONTENTS

II.A. DESCRIPTION OF RESOURCE OPERATIONS AND PROGRESS.

Our research resource aims to make the power of interactive computer graphics available to biohemists and crystallographers studying the conformation and function of proteins and nucleic acids. During the past year we have developed our prototype system to the level of functional completeness, convenience of use, and reliability necessary for productive use by chemists in original research. Since July 1975 we have operated on regular schedule and have served five groups of chemists, who have used a total of 545 hours of graphics system time.

II.A.1. Administrative Structure.

Our resource is based in the Department of Computer Science at the University of North Carolina. The attached chart shows the division of labor and relationships among those contributing to the resource this year.

Resource Facility Administration and Development

The Department of Computer Science's Computer Graphics Laboratory, the facility we use, was managed as a service facility by Mr. James Robb, the department's associate chairman. Under his supervision was Mr. Peter Nichols, an electrical engineer in charge of the equipment. Because funding for 75-76 was not as much as in 74-75, Mr. Robb resigned in December; Mr. Nichols was laid off. The facility is currently managed by Dr. J. D. Foley.

Software support for our graphic facility includes several locally developed programs: a S/360-PDP11 cross-compiler for a substantial subset of the PL/1 language, routines for communication between the PDP11 and its host S/360 computer, routines enabling the PDP11 to access disk files attached to its host, and graphics subroutine packages. This software has been developed and maintained by a group of research assistants under the direction of Dr. Foley.

of SH Kim
of J Jueman
of D Richardson
s J Richardson
(all Duke)
of N Lipscomb
(Harvard)
of C Carter
(UNC)
D Ferro
(UNC)

Prof F P Brooks —
PI

Prof W V Wright
(PhD Student)
System Service
and Enhancement
and
GRIP II Design

Mr E Britton
(PhD Student)
GRAB Development
and Documentation

Prof V L Wallace
Exploratory
Development

Prof J D Foley
Graphics Facility
Development
and
(Since Jan 1)
Graphics Facility
Operation

Mr J L Robb
(Assoc Chairman)
(to Dec 31)
Graphics Facility
Operation

— T Bentley
(Spring '75)
— J Crawford
— T Dineen
— R Motley
— W Sidell
— D Tolle
(Summers)

— J Lipscomb
3-D Visualization

— M Pique
Multiple Bond
Twisting

— P Cisneros
Secretary

— W Babich
PhD

— G Hamlin
CAGES

— R Herron
— C Dunham
— R Jones
DIES

— J Warner
Business

— P Nichols
Elec Engineer
(to Dec 31)

Prof J Jukema
Prof J McQuaie
REFINE

Prof SH Kim
Dr J Xixxxx

use of N/t
Crypt System
via Tektronix
4014

1. All persons not otherwise identified are students in the UNC Dept of CS
2. Not all persons shown here were compensated from the N... grants (or at all)
3. The UNC-CS Graphics Facility is a multi-purpose facility used for research projects as well as molecular graphics.

UNC Molecular Computer Graphics Resource

Composition 1975-76

3

Service Support

The facility is made available to our client chemists on
a regular schedule of fifteen hours per week. Additional
time is available for both production use by chemists and
for system development, in competition with the other users
of the facility. Allocation of the available graphics time
among the client chemists is managed by Dr. W.V. Wright.

Dr. Wright directs a team of computer science research
assistants who help chemists prepare data for the system,
tutor them in its use as needed, and accompany production
sessions as needed. A typical production session has a che-
mist at the controls as pilot, a collaborator by his side as
copilot, and when necessary, a computer science research
assistant handy as flight engineer.

Development of Molecular Graphics System

The original GRIP system for displaying molecules was
developed by Dr. Wright as a dissertation project under Dr.
Brooks. That system has been converted to the current equip-
ment and later developments have been integrated into it.

The team under Dr. Wright continues to do enhancements
to the GRIP-GRAB System, and maintains the official current
version used for service and for a development base. Dr.
Wright himself not only coordinates development of GRIP-GRAB
enhancements but also coordinates the design of GRIP-II.
Dr. Wright is working 4/5 time as an investigator on our
resource project through a joint study agreement between UNC
and his employer, IBM. This agreement was renewed for the
two-year period 1 Jan. 1976 - 31 Dec. 1977.

Mr. Edward Britton, a Ph.D. student under Dr. Brooks'
direction, working with research assistants M. Pique and J.
Lipscomb under Dr. Wright's direction, converted GRIP to
the Vector-General, converted its molecular manipulation
from character-string commands to real-time analog commands,
and added electron density maps and their contouring. This
major package called GRAB gave us the first production sys-
tem usable for crystallography. It went into service in
July 1975, and has been the basis for all further work.
This year Mr. Britton has been documenting GRAB as his doc-
toral dissertation.

Drs. Wright, Brooks, and Wallace have been working as
the design team for GRIP-II. In addition, Drs. Wallace and
Foley serve as consultants on human factors aspects and on
many questions concerning interactive computer graphics.

Exploratory Development.

Mr. Michael Pique, who worked on GRAB under Dr. Wright in 1974-75, has this year been developing a facility for twisting a substructure about multiple rotatable bonds. This is his M.S. thesis work, advised by Dr. Wallace. His facility is about to be integrated into the standard system. Mr. James Lipscomb, who also worked on GRAB under Dr. Wright in 74-75, is completing a M.S. thesis under Dr. Wallace on 3-D Visualization. His developments have now been integrated into the standard system.

Mr. Jon Bentley worked under Dr. Wright on improved near-neighbor algorithms based on trees. He showed that these do not achieve full power for (1) dense structures such as molecules, and/or (2) those of the size of the molecules we have used.

Dr. John McQueen, a post-doctoral researcher working under Dr. Jan Hermans, prepared a scaled-down version of REFINE for the on-line (45-60 seconds) idealization of molecular structures. This has been integrated into the standard system.

Users and Collaborators:

All of the system users also serve as consultants in the development of the system, suggesting new functions and needed improvements in convenience. They also help assign priorities to these items. Of these Dr. J.L. Sussman of Duke University has been particularly helpful.

Our client users at Duke University also use a Tektronix 4014 display unit furnished under this grant. This terminal has been used primarily for access to the NIH CRYST system. Software for this terminal has been adapted from available programs by Dr. Sussman in collaboration with Dr. Richard Feldman of NIH. WE plan to develop software enabling this terminal to access the UNC molecular graphics system.

Drs. Hermans and McQueen have continued the development of the full-scaled batch version of REFINE. This program is now being distributed to other chemists.

II.A.2. Scientific Achievements

Core Research and Development.

This is an enumeration of the more important additions that our resource personnel have made to the system during the past year. A brief description of each of these enhancements is given in section IV.A. There enhancements fall into two classes.

1) Interactive Functions

- Calculation of contour maps from given electron density data and their display.

- Real-time manual manipulation of selected parts of a molecule and posting of the final conformation to the model.

- Many aids for visualizing the models and maps in three dimensions.

- Idealization of molecular geometry in selected parts of the model

- Recording of the picture being displayed at any moment.

2) Bridges to other programs and off-line storage

- Routines for transmitting atom coordinates and density data between our system and the S/370 model 165 located at TUCC. (Our users execute most large programs from building and refining molecular data on the TUCC machine.)

- Routines to save molecular models and density data on tape and off-line disk packs at the UNC Computing Center and to recover their data.

- A routine to construct hard-copy plots of any picture displayed during an interactive session.

Support Software Development.

Work continued on application-independent facility software, as described above under II.A.1.

Collaborative Research.

1) REFINE The programs used in our system to idealize molecular geometry on-line were derived from the REFINE package, a set of FORTRAN programs developed by Dr. Jan Hermans and his group supported by a grant from the National Science Foundation. Dr. Hermans has been collaborating with Drs. Jensen and Watenpaugh at the University of Washington and has made his programs available to a number of other groups including Dr. Feldman at NIH.

2) Storage-Tube Graphics. A Tektronix 4014 Storage-tube Display was installed in the biochemistry laboratory at Duke. Dr. Sussman has collaborated with Dr. Richard Feldman of NIH to implement access to the CRYST system from this terminal. They have added to CRYST some functions needed for analysis of molecular structures by X-ray crystallography. The resulting facility has been useful for making pictures of electron density maps and molecular structures for study and publication.

Service

Since July 1975 we have made our system available to five groups of biochemists for the purpose of carrying out original research. Their use falls into two general classes: (1) use as an electronic Richards Box, and (2) the comparison of two or more similar molecular structures. The use of our system as a substitute for a Richards Box has ranged from an attempt at an *ab initio* fit (the map was not good enough), to refinement of a structure for which an initial fit was available, to the fitting of a complex of a well-understood structure. Detailed descriptions of these applications of the system are given in Sections IV.A. We facilitate comparison of molecules by giving the user real-time control of the relative intensities of their superimposed displayed structures.

Training

No significant use was made of the system for education of biochemistry students during the year.

II.A.3. Plans and Objectives.

Our short and long-range plans are set forth in gory detail (80 pages of technical description plus appendicies,

etc.) in our continuation proposal dated August 26, 1975, and numbered 2-P41-RR-00898-03. These plans have been slipped a year, but not changed substantially. We shall briefly review them here.

GRIP works; REFINE works; GRAB works surprisingly well. GRAB and REFINE are in routine productive use and have assisted in the production of substantial published chemical results.

Our plans fall into two parts.

1. We want to establish a lab, a display system, a dedicated small crew, and a stable software system as a Regional Production Resource, for collaborative use by biochemists and crystallographers from the whole Southeast (and anywhere else as capacity allows).

2. We have begun a ground-up redesign of a GRIP-II system combining GRIP and GRAB functions and using the lessons learned from actual use of the existing systems.

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 6 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 learn more quickly about the needs of biochemists and crystallographers.

This strategy has been quite productive for us. The clear consequence, however, is that we at some point had 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." Hence the GRIP II design.

It is organized to meet the objectives delineated in our original proposal:

* It will be product-quality software, well-documented, exportable, and usable by strangers.

* It will be usable from a variety of graphics display terminals, from a variety of minicomputers, and from a variety of main computers.

- Where a configuration includes a powerful satellite (mini) computer, it will process all user actions not requiring extensive computations. If powerful enough, it will be able to execute the whole system stand-alone.

- 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 computer, 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 to GRIP and GRAB, to add functions, and to simplify using procedures.

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 continue to be integrated into GRAB.

We will undertake to implement a scaled-down version of GRAB for use with the Tektronix display in Dr. Kim's lab.

Our facility work will continue. In particular, we plan to investigate still more methods for enhancing three-dimensional perception.

II.A.4. 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 exceed, however, those from any 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.

14 May 1976

II.B. SUMMARY OF COMPUTER RESOURCE USAGE

Since we share the UNC Computer Graphics Laboratory, rather than having an NIH-furnished resource, a description of its use seems more appropriate than the conventional form.
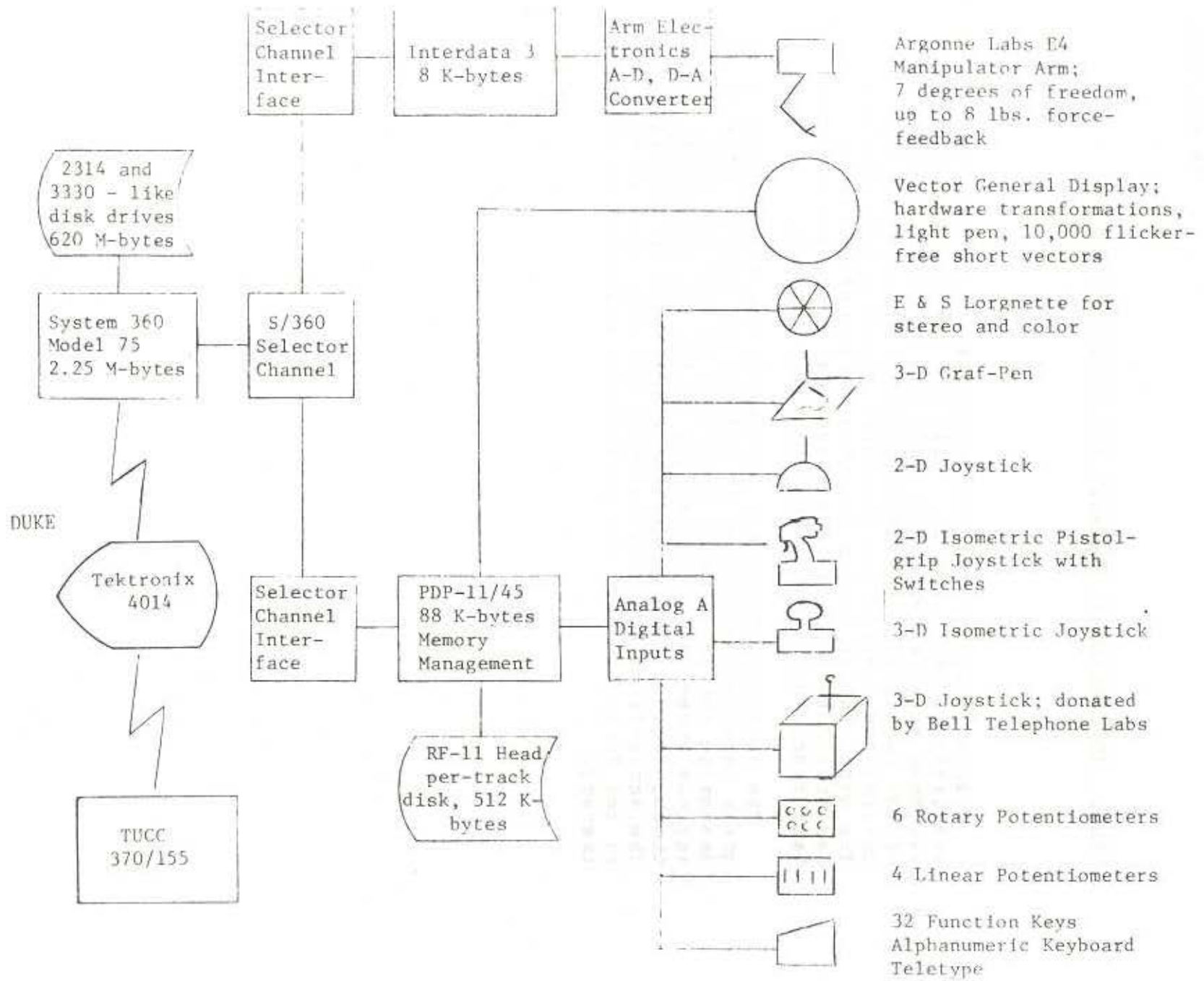
UNC Graphics Laboratory Usage 1975-76
15 July 1975- 30 April 1976

1. Molecular Graphics (all under NIH 5-P07 RR00898-02)

    Production Use   545 hours
      Kim - Sussman
      D. & J. Richardson
      Carter
      Ferro
      Lipscomb

    Development   342

    Demonstration   51
                                                     ‾‾‾‾
                                                     938

2. Other Research Projects (non health-related)   170

3. Education (non health-related)   453

4. Facility and Software Development   151

5. Maintenance of hardware   68
                                                    ‾‾‾‾‾
                                                    1780

II.C. THE RESOURCE FACILITY AND EQUIPMENT LIST

Figure 2 depicts the entire complement of equipment available to our resource. Only the Tektronix 4014 at Duke has been acquired with NIH funds and is solely used as part of the resource. The display console and most of the interaction devices are clustered around a work station. The work station design is our second complete design. It has proved to be much more convenient than that reported last year. We do not contemplate any changes.

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.

Resource Diagram

Figure 2

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

System 360 Model 75 2.25 M-bytes — S/360 Selector Channel

Vector General Display; hardware transformations, light pen, 10,000 flicker-free short vectors

E & S Lorgnette for stereo and color

3-D Graf-Pen

2-D Joystick

DUKE

Tektronix 4014

Selector Channel Interface — PDP-11/45 88 K-bytes Memory Management — Analog A Digital Inputs

2-D Isometric Pistol-grip Joystick with Switches

3-D Isometric Joystick

RF-11 Head per-track disk, 512 K-bytes

3-D Joystick; donated by Bell Telephone Labs

TUCC 370/155

6 Rotary Potentiometers

4 Linear Potentiometers

32 Function Keys Alphanumeric Keyboard Teletype

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	PXYT	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

14

RESOURCE EQUIPMENT LIST

Equipment Located Outside the Main Resource Area

Location/ Description/Identification	Equipment			Date Installed	Date Accepted	Cost		
	Manufacturer	Type	Model			Purchase Price	Annual Rent	Source of Funds
Storage Tube Display	Tektronix		4014	4/30/75	4/30/75	13,000		NIH

LOCATION: Duke University, Durham, North Carolina

II.D. SUMMARY OF PUBLICATIONS

Carter, C. W., Jr. "X-ray Analysis of High Potential Iron Protein and Ferredoxine" in IRON-SULFUR PROTEINS, Academic Press. Editor, Walter Lovenberg. In Press

Foley, J. D. and Wright, W. V. "An Interactive Molecular Graphics System with a Satellite Terminal Closely Coupled to its Host", (Panel discussion position paper). Proc. ACM 1975 Annual Conference: 88-89.

Hermans, J. et al "Use of Energy Calculations to Refine and Understand Structure and Function of Proteins". Proc. Symposium on Environmental Effects on Structure and Properties, Jerusalem, 1975.

Kim, S. H., "Symmetry Recognition Hypothesis Model for tRNA Binding to Aminoacyl-tRNA Synthetase". Nature 256, 679, 1975.

Kim, S. H., Sussman, J. L., and Church, G. M. "A Model for Recognition Scheme between Double-stranded DNA and Proteins". In Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions (Sundaralingam and Rao, Eds.), University Park Press, Baltimore, p. 571, 1975.

Richardson, J. S.; Richardson, D. C.; Thomas, K. A.; Silverton, E. W.; and Davies, D. R. "Similarity of Three Dimensional Structure Between the Immuno-Globulin Domain and the Cu, Zn Superoxide Dismutase Subunit, Journal Mol. Biol., In Press.

Kim, S. H., and Sussman, J. L. " Turn in a Conformational Pattern in RNA Loops and Bends". Nature Vol 260, No. 5552, p. 645. 1976.

Sussman, J. L. and Kim, S. H. "Idealized Atomic Coordinates of Yeast Phenylalanine Transfer RNA". Biochem. Biophys. Res. Comm. 68, 89, 1976.

Sussman, J. L., and Kim, S. H. "Three-Dimensional Structure of a Transfer RNA Common in Two Crystal Forms". Submitted for publication in Science.

III.A. SUMMARY OF RESOURCE EXPENDITURES

| | | BRB Allocation | | Total Resource Expenditures | |
| | | Actual | | Actual | |
		Previous Budget Period	Current Budget Period	Previous Budget Period	Current Budget Period
1.	Personnel				
	a. Salaries & Wages	46,321.00	55,192.00	111,441.00	94,577.00
	b. Fringe Benefits	1,890.00	4,927.00	3,631.00	5,975.00
	Subtotal	48,211.00	60,119.00	115,072.00	100,552.00
2.	Consultant Services	708.00		2,514.00	
3.	Equipment				
	a. Main Resource – Rented		1,100.00	1,978.00	3,320.00
	b. Main Resource – Purchased	172.00	985.00	1,650.00	1,734.00
	c. Supporting Equipment				
	d. Equipment Maintenance		2,500.00	2,903.00	2,624.00
	Subtotal	172.00	3,385.00	6,531.00	7,678.00
4.	Supplies	1,090.00	799.00	7,313.00	2,187.00
5.	Travel	1,490.00	318.00	2,048.00	628.00
6.	Alterations & Renovations				
7.	Publication Costs			397.00	213.00
8.	Other				
	a. Computer Services		3,000.00	15,965.00	7,450.00
	b. Other	17,405.00	1,165.00	17,405.00	1,450.00
	Subtotal	17,405.00	5,365.00	33,370.00	8,900.00
9.	Subtotal – Direct Costs	69,076.00	69,986.00	167,245.00	120.158.00
10.	Indirect Costs				
	a. Previous				
	b. Current @ 20%	13,817.00	26,564.00	35,411.00	31,362.00
	c. Next @ 20%				
11.	Total Cost	82,900.00	96,550.00	202,665.00	151,520.00

III.B. EXPENDITURE DETAILS
Direct Costs Only

Current Budget Period

		% of Salary from BR Grant	Amount	% of Time or Effort	(Hours) Amount
1.	**PERSONNEL**				
	1. Professors				
	F.P. Brooks, Jr.	-	-	7%	145
	J. Hermans	12%	3,600.00	10%	208
	2. Associate Professor				
	V.L. Wallace	31%	6,715.00	25%	519
	3. Assistant Professor				
	J.D. Foley	8%	1,889.00	8%	166
	4. Pesearch Associate				
	P. Nichols	83%	5,872.00	50%	1038
	J. McQueen	100%	6,996.00	100%	2076
	5. Secretary				
	P. Cisneros	100%	7,080.00	100%	40
	6. Research Assistants				
	E. Britton	29%	427.00	100%	20
	J. Crawford	17%	670.00	50%	10
	T. Dineen	94%	3,626.00	100%	20
	R. Hogan	68%	1,920.00	68%	13
	M. Ling	100%	1,920.00	100%	20
	R. Motley	100%	3,849.00	100%	20
	M. Pique	100%	1,920.00	100%	20
	J. Walker	50%	1,920.00	100%	20
	J. Warner	42%	1,242.00	100%	20
	Fringe Benefits		4,947.00		
	Total		54,597.00		

EXPENDITURE DETAILS (continued)

		Current Budget Period	
		BR	Total
2.	Consultant Services	–	–
3.	Permanent Equipment		
	Main Resource - Rented		
	1. IBM Copier I	1,087.00	1,087.50
	2. IBM Copier I	–	720.00
	3. 2314 Disk Pack	–	1,500.00
	4. Other	12.50	12.50
	Subtotal	1,100.00	3,320.00
	Main Resource - Purchased		
	1. Vector General FCO (and other parts for a modification of the VG system).	–	260.00
	2. Estimated to be spent by 6/29/76	985.00	985.00
	3. Other	–	498.00
	Subtotal	985.00	1,734.00
	Equipment Maintenance		
	1. Repair of Interface Control	–	124.00
	2. Digital Equipment Corp. Maintenance Contract	2,481.00	2,481.00
	3. Other	18.20	18.20
	Subtotal	2,500.00	2,624.00
4.	Consumable Supplies		
	1. Electronic parts	15.30	15.30
	2. Electronic parts	–	791.00
	3. Film	142.00	142.00
	4. Paper	60.00	60.00
	5. Paper	–	575.00
	6. Printed Information	59.00	59.00
	7. Printed Information	–	22.00
	8. Computer Cards	10.00	10.00
	9. Other	512.70	512.70
	Subtotal	799.00	2,187.00

EXPENDITURE DETAILS (continued)

	Current Budget Period	
	BR	Total
5. Travel		
1. Trips to Conferences	318.00	318.00
Total Travel		310.00
Subtotal	318.00	628.00
6. Alterations and Renovations		
None		
7. Publication Costs		
1. Estimated to be spent by 6/30/76		213.00
Total Publication Costs		213.00
8. Computer Services		
1. UNC Model 360/75		
2. Varying Rates		
Total Computer Services	3,000.00	3,000.00
		4,450.00
Subtotal	3,000.00	7,450.00
9. Other		
1. Telephone, Postage, Packing, Shipping		
Total Other Expenditures	1,165.00	1,165.00
		285.00
	1,165.00	1,450.00
GRAND TOTAL	9,867.00	19,606.00

III.C SUMMARY OF RESOURCE FUNDING

	BUDGET PERIODS	
	Actual Previous Budget Period	Current Budget Period
SOURCE OF FUNDS		
Service Charges	None	None
Institution Funds	None	None
Other Support Grant Funds		47,640.00
BR Support Amount of Current Award Direct Costs Indirect Costs Total		96,550.00 69,986.00 26,564.00 96,550.00

New Air Force grant is not in support of resource.

IV.A.1. The Interactive System.

Our current interactive computer graphics system for molecular studies is primarily the work of Messrs. Britton, Lipscomb, and Pique all of whom are presently completing the requirements for degrees in computer science at UNC. These three, working under Drs. Wright and Brooks, based their design on and converted programs from the pilot system for molecular graphics built by Dr. Wright in 1970. This work has been continued during the past year by three first-year graduate students, Mrs. Motley and Messrs. Dineen and Tolle. The most important interactive functions of this system are:

a. Display of Kendrew-type stick-figure models of proteins and nucleic-acid structures.

b. Real-time joystick control of the direction from which the molecule is viewed. The user can smoothly turn the picture. This gives an imitation of head-motion parallax, an important depth cue.

c. Stereoscopic display of molecules using an Evans & Sutherland lorgnette, depth cueing by intensity, depth cueing by a Z-axis clipping plane, depth cueing by steady rotation or oscillation, and stereoscopic display by half-silvered mirror and polarized filters (over-under stereo). 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 much better than a picture presented using only one of the 3-D techniques.

   This work has been done by J. Lipscomb under Dr. Wallace's direction. Mr. Lipscomb has not been supported by grant funds.

d. Selective display of individual residues in a molecule, and control of the detail displayed. Because a picture showing all bonds of the entire molecule is confusing for interesting macromolecules, a facility has been added for displaying three levels of detail: (1) a set of lines not corresponding to

21

valence bonds but connecting the principal atoms of
each residue (e.g. the alpha carbon atoms of a pro-
tein); (2) all bonds of the main chain, or (3) all
bonds. These levels of detail are illustrated by
pictures of the phenylalanine transfer RNA (tRNA)
molecule in Figures 3 through 6 The level of detail
displayed can be controlled independently for each
residue. This can be seen in Figure 5 in which all
bonds in a selected region of the tRNA molecule are
shown and in Figure 6 which shows part of this same
region expanded (and rotated) so that it fills the
display screen.

e. Manual analog specification of a point in the model
space. A joystick with three rectilinear degrees of
freedom is used to move a cursor about in the dis-
play screen. This enables the user to specify a
point in the model space more quickly and more easi-
ly than with the keyboard entry and light-button
commands of the original GRIP system. Such input is
useful for moving a molecule to an approximate posi-
tion. The directions of cursor movement are those
of laboratory space.

f. Interactive contouring and display of electron
density maps. These density maps are displayed as
contours of equal density drawn on a set of parallel
planes evenly spaced through the visible portion of
the model space. Figure 7 shows the portion of the
tRNA model illustrated in the previous figure
embedded in the corresponding electron density map.
The fitting of this model to its density map as
shown in this figure has not been completed, and a
number of discrepancies are apparent in this pic-
ture. The user specifies the density levels to be
shown, and several selected densities can be shown
simultaneously. The contour planes are always per-
pendicular to one of the coordinate axes of the
model space. The user can choose any one, two, or
all three of the possible sets of orthogonal planes
for simultaneous display. Figure 8 shows the region
of the previous figure with two sets of contour
planes displayed. Another feature allows the user
to suppress the display of any individual contour
and, if desired, to reinstate it later without
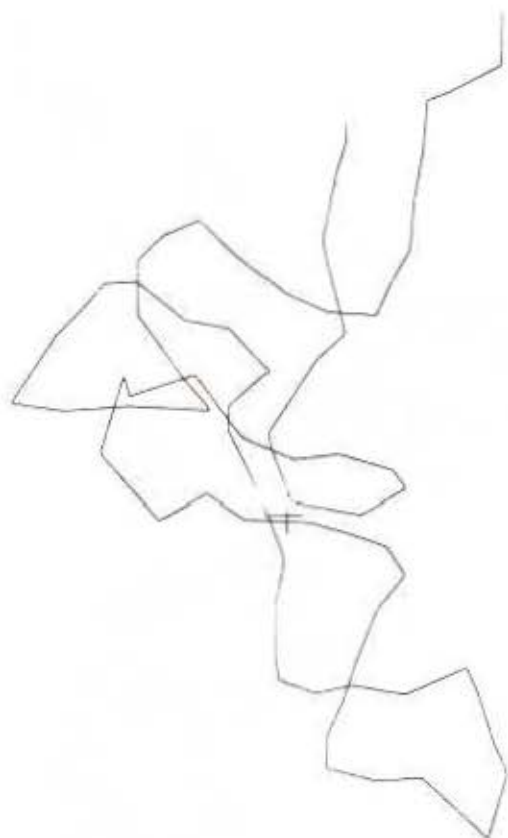recalculation.

22

Figure 3   Backbone of tRNA-- the C_1 atoms of adjacent residues
                  in the molecular sequence are connected by line segments.
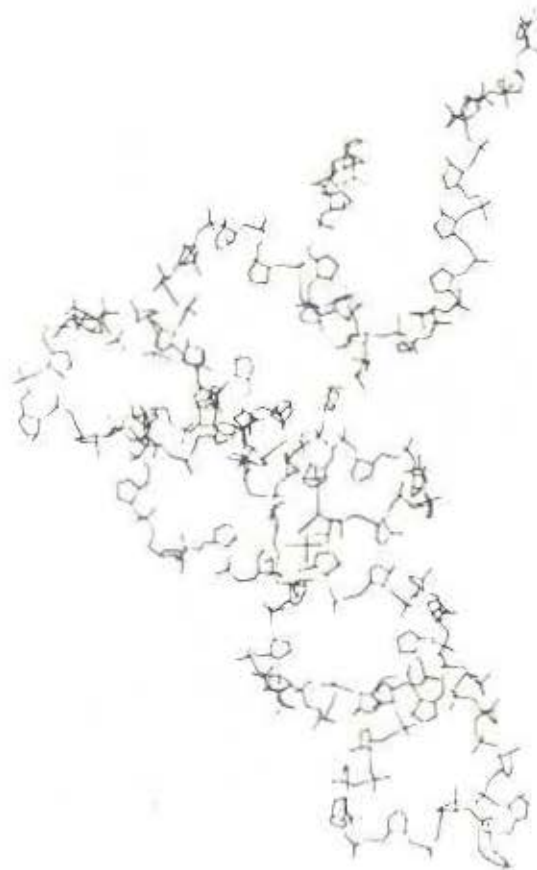
Figure 4   Main chain of tRNA-- each valence bond in the main chain
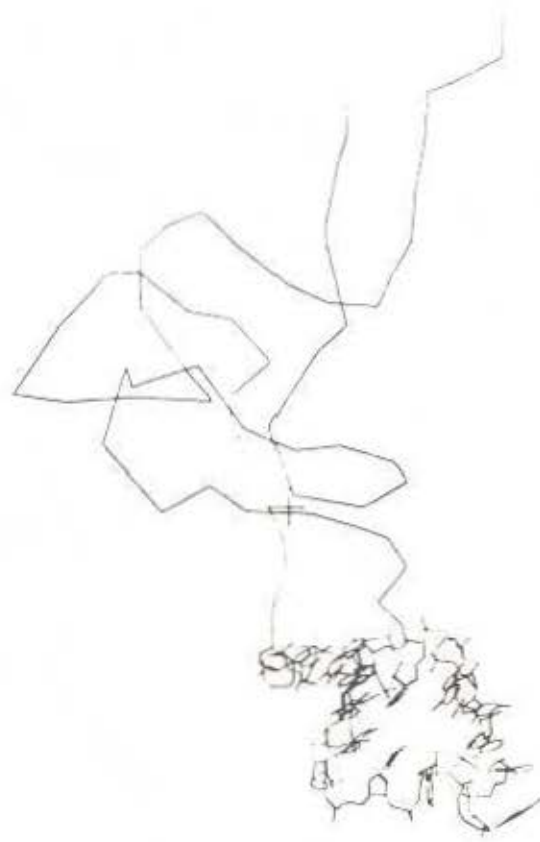   is represented by a line segment.

Figure 5   Selected detail of tRNA-- all valence bonds of residues
                  28 through 42 are shown with the backbone of the remainder
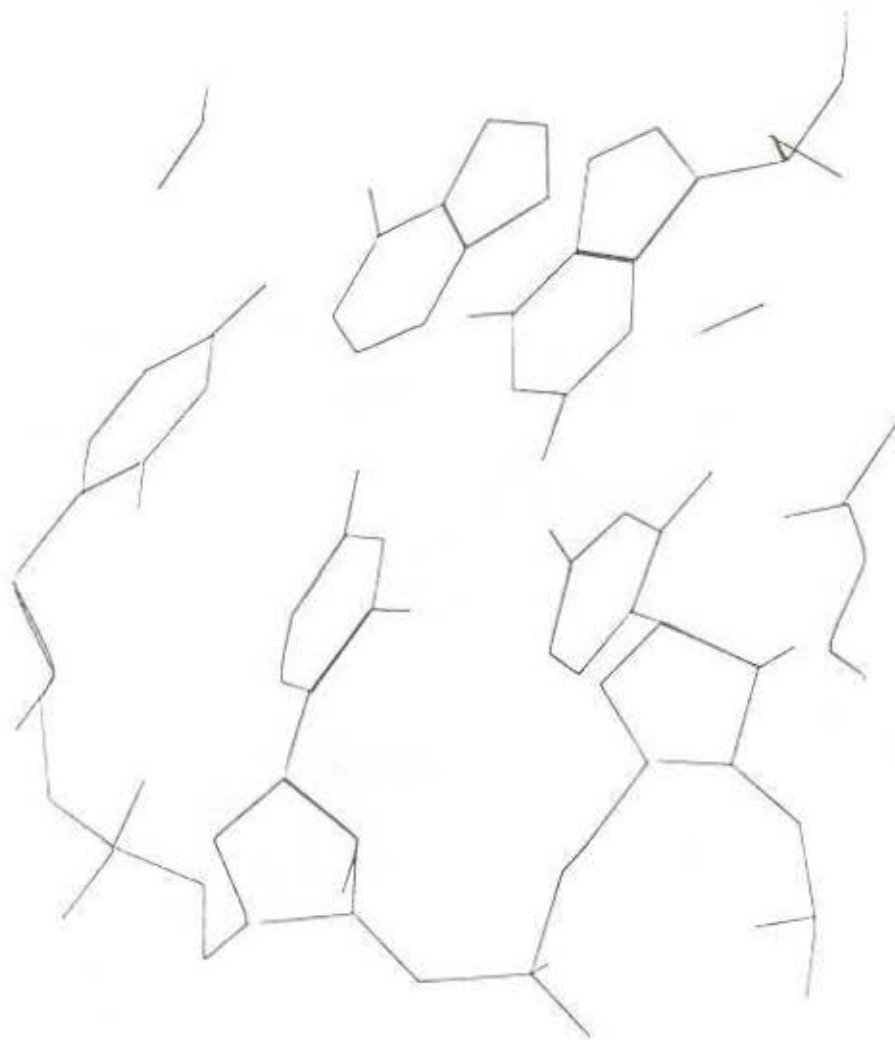   of the molecule.

Figure 6   Enlarged detail of tRNA-- a portion of the previous figure
                  enlarged six times and rotated.
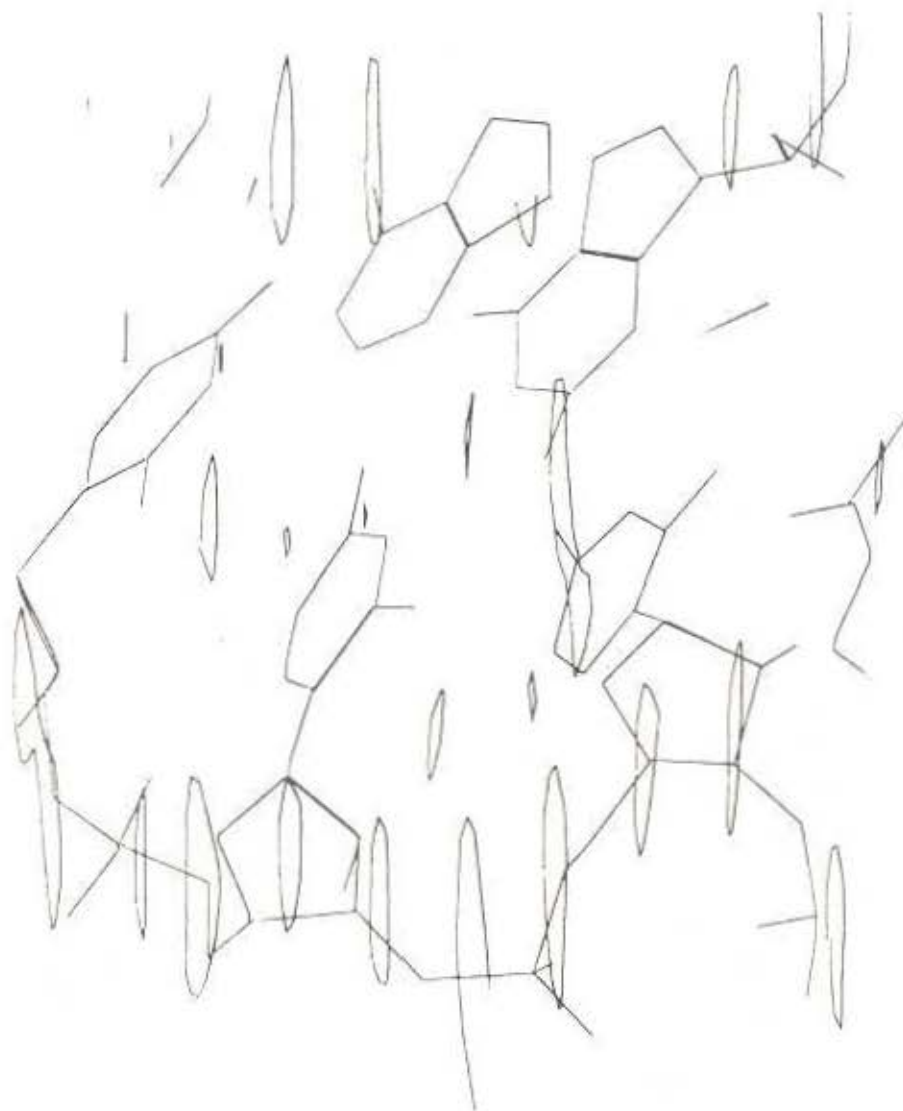
Figure 7   tRNA model and electron density map-- the portion of tRNA
            shown in the previous figure embedded in its electron
            density map which is shown as a set of contour maps on
            parallel planes.

q. Manual fitting of a model and electron density map in real time. The user can select a part of his model (e.g. a residue) and can manipulate its position and orientation by means of two separate manual input devices with three degrees of freedom each. He can also twist up to eight selected bonds in the selected substructure by turning one knob at the workstation for each bond. The new position, orientation and conformation of the selected substructure are dynamically shown and can be recorded in the molecular data structure upon command. This is an essential facility for fitting a model to an electron density map or any other visually determined conformation. This facility was developed by Mr. Pique working under Dr. Wallace's direction.

h. Plotter output. The user can record the picture he sees on the graphic screen at any time. A complete description of this current picture is stored on disk where it can be used to generate a permanent copy via a plotter or a line printer. This facility was used to generate Figures 3 through 8 of this report. It also provides a means for making large reproductions of an electron density map or a molecular model.

i. Real-time control of the relative brightness of the map and model on the display screen (or of two models if so map is displayed). This helps the user distinguish his map and model when fitting and is useful for comparing.

j. Interactive computational refinement of a selected part of the model to minimize deviations from the ideal bond lengths and angles. This function, built by Dr. McQueen, uses an on-line version of the REFINE program developed by Dr. Hermans and his group.

k. Library management. The GRIP system now includes several libraries of molecular models recorded on disk. Commands are provided for displaying a list of the models in any selected library, for loading a model from a library into the core memory of the system, for saving the current model in a library, and for deleting a model from a library. It is intended that each user should have his own library of models. Models can also be entered into the sys-

tem from punched cards or magnetic tapes. The cur-
rent version of this facility was built by Mrs.
Motley.

1. Recovery from a system failure. A facility is pro-
    vided which records the current molecular model on a
    disk library after every user command which modifies
    this model. Thus, the most recent molecular model
    is always available on disk. When the system is
    restarted after a normal or abnormal termination,
    this model can be reloaded into core memory. These
    backup models are recorded alternately in two separ-
    ate libraries to give some protection against a sys-
    tem failure during the recording process. In the
    near future we plan to add facilities for recording
    the system control information which specifies the
    current view of the molecules in addition to the
    model data.

Human Factors. We have devoted a considerable part of our
effort to building a system with smooth human factors. For
example, we have selected manual input devices whose degrees
of freedom are intuitively related to the model parameter
control. Also, all manual controls except bond-twisting
work in laboratory space instead of model space.

Releases. As we began to use the systems for productive
research it became obvious that evolving experimental sys-
tems are not suitable for productive work. Therefore, we
have constructed a sequence of stable, consolidated releases
of our system for production use. Each of these inco-
rporates all of the facitilies proven workable by the chosen
cut-off date. The most recent release represents our status
as of February, 1976.

IV.A.2. Bridges to batch programs.

We have begun to build bridge programs between GRIP and a
number of existing batch programs for building molecular
models from residue sequence data, for refining structures,
and for calculating electron density maps.
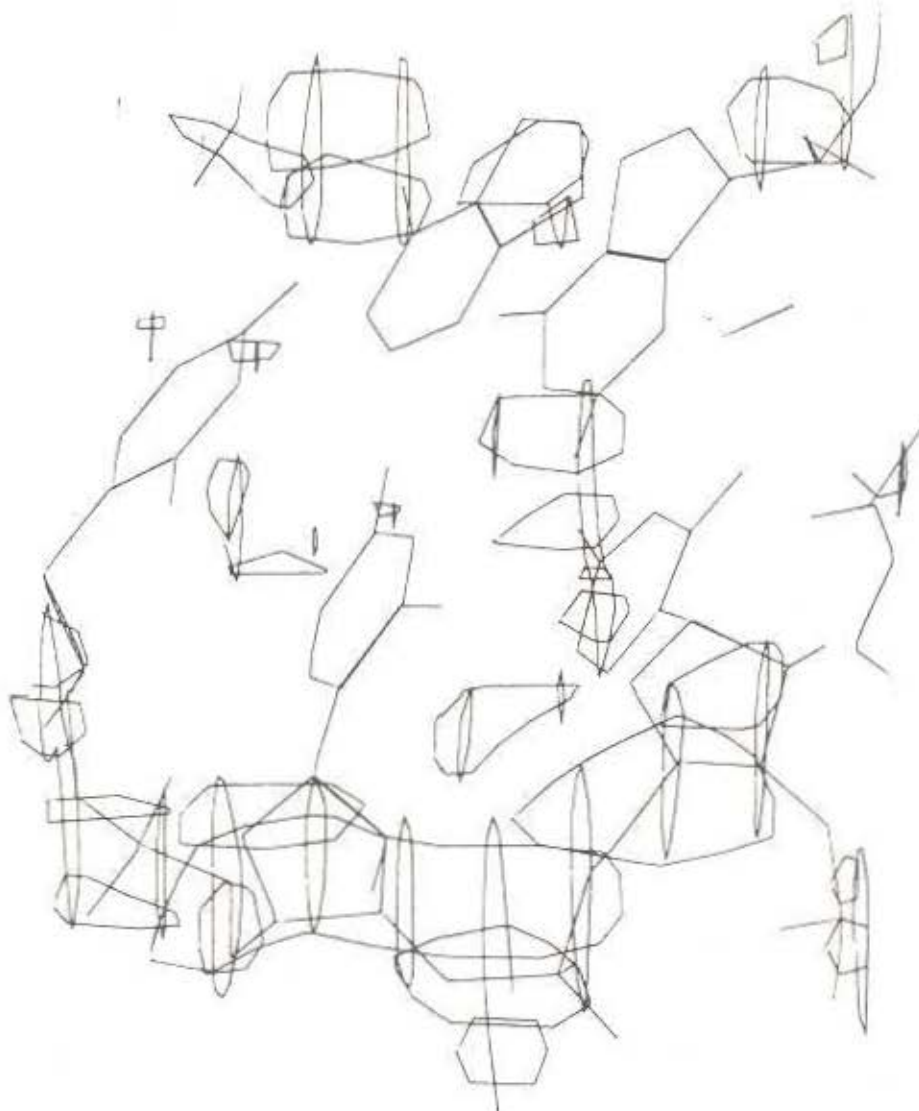
Figure 8   Two sets of orthogonal contour maps-- the previous
                figure with another set of contour maps added which are
                drawn on planes perpendicular to those used for the first
                set of contour maps.

IV.A.3. Configurability- CAGES

We intend several basically different modes of operation:
closely coupled, nearly stand-alone, and stand-alone. There
are also many modes in between the extremes. Our objective
is to produce software which can be very easily adapted to
any of the modes. We now have in hand the basic tools
needed to do this: the first tool is compatible compilers
for the 360/75 and 11/45. Our first such was developed
before this grant started; a second-generation PLCD compiler
is now under development. The second has been completed and
documented this year, a system called CAGES - Configurable
Applications for Graphics Employing Satellites, written as
part of G. Hamlin's dissertation work.

CAGES is a programming system which substantially simpli-
fies the process of writing interactive graphics application
programs for use in a distributed processing, satellite-host
configuration. It allows programs written in a PL/I subset
to be configurable; program modules and data can be easily
reassigned from the host to the satellite, or vice versa.
That is, the division of labor between the two computers is
readily modified.

The CAGES system supplements the operating system ser-
vices normally provided on the host and satellite computers
by providing the illusion that the application is executing
in a single computer memory with dual CPU's. In reality the
application is distributed between the memories of the host
and satellite computers. To maintain this illusion CAGES
provides three types of services:

a. Inter-computer subroutine CALLs and RETURNs. A
    subroutine call executed on one computer and tar-
    geted to a subroutine resident on the other com-
    puter is known as a remote procedure call. A mes-
    sage containing the subroutine's name and parame-
    ters is sent to the computer where the subroutine
    is to be executed. Following its execution, a
    message containing results is sent back, and the
    calling routine continues its execution.

b. Inter-computer signals. A SIGNAL statement or
    interrupt on one computer can raise a CONDITION
    for which there is an enabled ON-BLOCK in the oth-
    er computer.

c. GLOBAL data references. Variables with EXTERNAL
    scope which may be referenced from subroutines in
    both computers are known as GLOBAL variables. The
    CAGES system handles such references by obtaining
    the needed data from the remote computer and plac-
    ing a local copy of it in the memory of the
    referencing computer. The program is then allowed
    to access this copy.

Beyond the mere capability of configuring an application
to a particular host-satellite system, CAGES allows the app-
lication program designer to view his program as a whole, as
if it were to be executed on just one computer, without
knowing its ultimate configuration. This is very helpful in
producing a conceptually integrated design. On the other
hand, ultimate efficiency of operation demands that the
designer recognize that the program will in fact be distri-
buted. To aid in designing optimally for distributed confi-
gurations, Hamlin has set forth a series of design and pro-
gramming guidelines. They are concerned with how the appli-
cation is modularized, how data is accessed, and inter-
module and intra-module patterns of reference to data files,
GLOBAL variables, and procedures. The guidelines have been
proven sound, both experimentally and theoretically. We
will use them.

Another advantage of CAGES is that it allows the division
of processing and data between the host and satellite to be
fine-tuned to maximize responsiveness. In this regard, we
will draw on the tools and insights developed by Foley's and
Wallace's Air Force-sponsored work on modelling of graphics
systems. Their goal has been to develop models and guide-
lines to be used in dividing processing and data between the
host and satellite.

IV.A.4. Portability- DIGS

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 to
our project's goal of making our programs usable from a
variety of display terminals, ranging from low-cost, low-
performance systems such as 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 displays with the exact set of options we have. A display-independent graphics package, DIGS, 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 [Sproull73], GPGS [Caruthers72], and GINO [Sancha73]. We have examined these, and others, in detail. Unfortunately, none of them meets the 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, augmented, or dynamically transformed. A viewport is a collection of segments. It too can be deleted, augmented (by adding segments), and dynamically transformed. Thus a two-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 (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 S/360 or S/370, the third for a 4014 coupled to a 11/45 and that in turn to a S/360 or S/370. The DIGS architecture has been completed [Foley75], and much of the internal organization has been designed. Some coding

for the 11/45-VG version has been completed. We intend to
use DIGS in GRIP-II.

IV.A.5. Near-neighbor algorithms.

Mr. J. Bentley, a fellowship student in the Department
of Computer Science, has been investigating algorithms suit-
able for finding all atoms in a specified neighborhood of a
molecular model. These algorithms have excellent asymptotic
performance as the number of atoms in the structure
increases. Bentley has concluded, however, that for dense
molecules of less than 2000 atoms, the partitioning-type
algorithms are superior to tree algorithms.

14 May 1976

IV.B.1. The REFINE Program

REFINE is a program package developed by Hermans in collaboration with 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:

    a.   test the stereochemistry of the coordinates.
    b.   build a model best fitting the coordinates.
    c.   build a model of low packing (nonbonded) energy.
    d.   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, to Feldman at N.I.H., to Meyer at Texas A & M, and Koetzle at Brookhaven. A third version has been completed and is being used by Hermans at U.N.C. A publication describing results obtained with model building calculations has appeared. [Hermans75].

We prepared a CDC version of REFINE2 and this has been installed at Brookhaven. The staff at Brookhaven has made the program accessible on the CRYSNET system (June 1975).

The CDC version of REFINE2 has also been installed at the University of Groningen (July 1975). It will soon replace the old version in use at the University of Washington. Feldmann and Sussman have integrated REFINE2 into the molecular graphics system at N.I.H.; this system is accessible via telephone lines. Sussman uses it at Duke with the Tektronix terminal. Program and manual are available to other groups on request.

Use of REFINE3. 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 excel-
lent geometry is obtained with energy minimization of a
strained ring structure. The second, on antamanide, a cycl-
ic sodium binding decapeptide, showed that energy minimiza-
tion of a preliminary X-ray structure gives improvement
towards the structure obtained with classical crystallo-
graphic refinement. These studies have been completed and
the results are available. [Hermans et al75]

In the third project we have almost completed a study of
the refinement of a much larger structure, that of the pro-
tein rubredoxin. (This work is being carried out in colla-
boration with Jensen's group, which determined and refined
this structure). As a fourth project, we have begun a study
of the activated state of the enzyme-substrate complex of
the cystine protease papain, using X-ray crystallographic
data obtained by Drenth's group in Groningen. The mechanism
of catalysis by papain is similar to that by the serine pro-
teases subtilisin, chymotrypsin, trypsin, thrombin, etc.

IV.B.2. Storage-Tube Graphics at Duke

There is a need for at least a primitive computer gra-
phics facility in the hands, i.e. the laboratory, of mole-
cular biologists to perform routinely the following
functions:

   a.   to visualize in detail and in three-dimensions
        macromolecular structures.
   b.   to alter the conformations of these structures.
   c.   to attempt to fit a three-dimensional model of a
        macromolecule to an electron density map derived
        from X-ray diffraction data.

As part of our initial grant, we acquired a Tektronix
4014 direct-view storage tube and installed it in Dr. Kim's
laboratory. Drs. Kim and Sussman have been exploring its
uses of the above purposes.

As several laboratories have developed programs to do
many of these operations already, we have worked with other
groups in extending and adapting their systems. So far the
bulk of our collaboratiion has been with Richard Feldmann of

34

the NIH in extending the CRYST system for our particular
needs.

The facility to display macromolecular structures (or for
that matter any structure in the Cambridge Crystal File) in
up/down polarized-filter stereo on a storage tube terminal
has been implemented for about two years on the NIH CRYST
system.

The CRYST system for molecular display has been a plea-
sure to use, as it is almost transparent to even a novice
user although extremely powerful. It is extremely well
documented and everyone in our laboratory who has tried to
use the system has become more or less expert in less than
an hour of real time use.

In our laboratory we have used the visualization features
of the CRYST system for two specific applications:

1)   Comparison of different macromolecular structures

Jane Richardson compared in detail the three-dimensional
structure of Superoxide Dismutase (SOD) whose crystal struc-
ture was recently determined in our laboratory [Richardson,
et al75] with several different immunoglobin structures
(solved in other laboratories) that appeared to have similar
structure although quite different function. The coor-
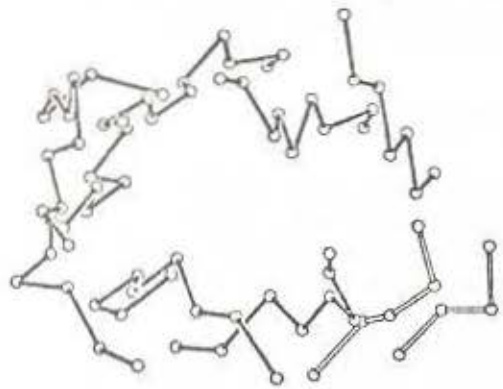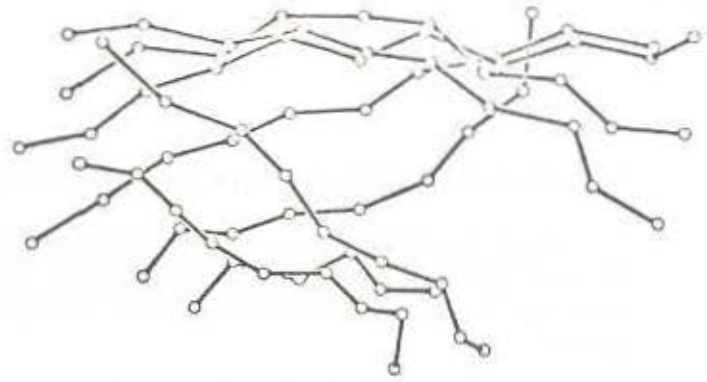dinates for these structures were already on the CRYST
system.

Figure 9 was prepared by redrawing the hand tracing of
the SOD and two other portions of a protein (V & C) from
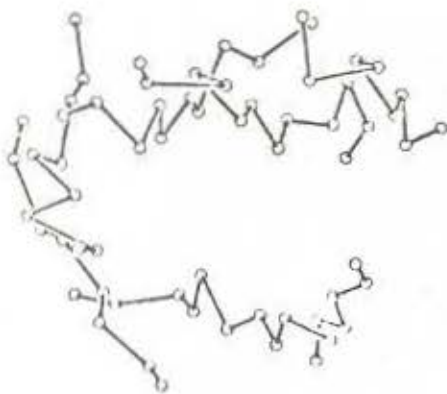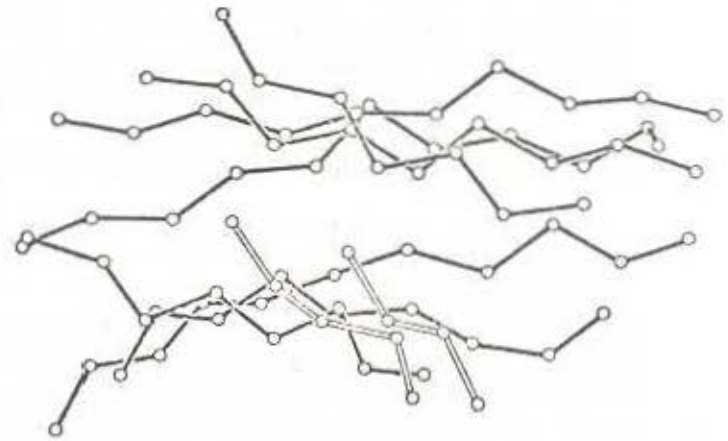two different points of view.
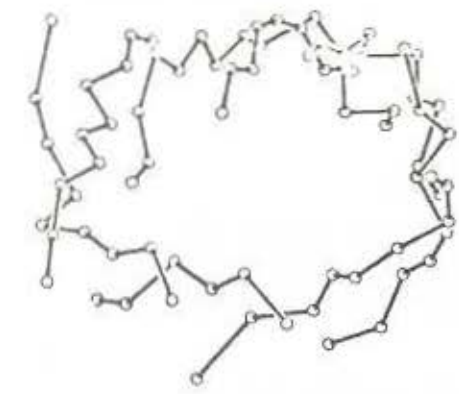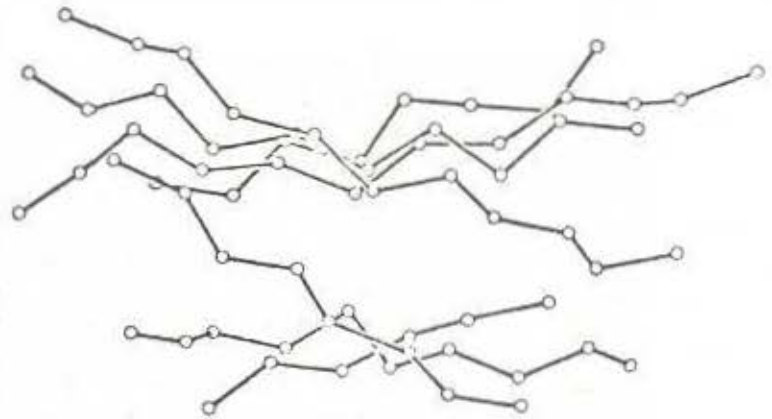
SOD

V_H

C_{H1}

Figure 9

Redrawn hand-tracings

2) Examination of the tRNA structure detail

Joel Sussman has used the CRYST system to display the three-dimensional structure of yeast-tRNA which was recently solved in our laboratory in order to scrutinize our current model.

So as to minimize telephone and computer costs we have used the audio cassette recorder option of our terminal in order to store various views and fragments of the tRNA structure on cassettes. In this way we have begun to build up a library of cassettes which can be used stand-alone on our terminal to view the tRNA structure.

3) Manipulation of macromolecular structures

In the course of a macromolecular structure refinement it is essential to know if one's model has reasonable stereo-chemistry and, if not, to have the facility to correct any shortcomings. The REFINE program performs just such a function. Over the past year, Sussman in our laboratory has modified this program to make it more suitable for nucleic acid structures and has used in on the TUCC computer without graphics capability in a batch mode.

4) Fitting a three-dimensional model to an electron density map

Sussman and Feldmann have adapted an electron density contour map drawing program (which Feldmann had written several years ago) to run with the storage tube as an output device. In addition they have "married" the molecular model to the map with all the molecular manipulation features of the CRYST system. Both the map and the molecule can be drawn in stereo. Generally one examines a portion of the map/model space of one to three residues, with the computer keeping track of all the bookeeping such as what part of the map to include, scale, center, etc. Thus they have developed a prototype Electronic Richards' Box which runs on a time-shared computer with a storage tube as its display vehicle.

The contour drawing program can also be used by itself to draw the entire map plane by plane. These sections can be copied onto transparent sheets of acetate (via a inexpensive thermal copier) and a mini map of the entire structure can be assembled quite quickly.

36

IV. C. CHEMISTS' USE AND CHEMICAL RESULTS

The system has been used on five biochemical research projects. Each of these projects represents a different stage in the investigation of a molecular structure.

A group working under Dr. W. Lipscomb at Harvard used GRIP to sketch a model of ACTase. This investigation is at a very early stage, with a low-resolution map.

Dr. David Richardson and Jane Richardson of Duke used GRIP to construct a model of Superoxide Dismutase. The Richardsons had preliminary coordinates for their alpha carbon atoms and are thus somewhat farther along than the Harvard group. This is the first instance we know of in which a graphics model was constructed without a Kendrew model being built beforehand or simultaneously.

Drs. Kim and Sussman of Duke are currently using the system to improve an existing model of tRNA. Their investigation has advanced considerably further than the two mentioned above. Refined coordinates have been published.

Dr. Ferro, in Dr. Hermans' laboratory at UNC, used the system to compare several alternative models of Rubredoxin which are already quite refined. His work thus represents the farthest advanced investigation for which GRIP has been used.

Dr. Carter of UNC used the system to compare the oxidized and reduced states of HIPIP.

Duke Users' Experience

We (Sung-Hou Kim and Joel Sussman) have used the UNC graphic display system in production as an "electronic Richards' box" for the purpose of fitting the crystal structure of yeast-tRNA to its electron density map.

We have found it to be far superior to a conventional Richards' Box in visualization of both the model and the map, as well as in the actual fitting by interactive manipulation of the model to the map. In addition we have found it much easier for two people to work together in trying to fit the model to the map than ever was possible with the conventional Richards' box.

Visualization. The visualization is better in that it is possible to get an optimal view of a particular region of the map/model space in order to do the fitting. This generally involves:

(1) Selecting a small part, a few residues, of the model and map space.

(2) Rotating the fragment of the model/map space until one sees just the right viewpoint to begin fitting. Generally this has been done with one-dimensional (planar) contours, since the lognette stereo limits the number of vectors that can be drawn. At present, one can have stereo plus one-dimensional contouring or monocular viewing plus two- or three-dimensional contours. If more drawing speed were available, we would routinely view density contours in two orthogonal planes (the basket-weave contours developed by Barry). This effect greatly enhances our ability to understand how a residue should fit into the map.

(3) Changing the contouring level(s) for the particular fragment of map/model space until the right contours are present to do the fitting without cluttering the picture with too many contours. We usually use one contour level at a time.

Fitting. The actual fitting generally proceeds in three steps:

a) FITRES - a residue (one nucleotide) is 'detached' from the structure, giving it 6 degrees of freedom, 3 rotations and 3 translations. We then attempt to fit the sugar and base portion of the residue to the electron density map. Once they fit we reconnect the residue to the rest of the structure (Freeze it).

b) FITPHOS - next the phosphate group for the same residue is 'detached' from the structure and again with 6 degrees of freedom it is fitted to the map, and frozen.

c) BONDTWIST - finally when a particular residue is involved in hydrogen bonding base pairing with another residue in the structure, both residues are displayed, and by means of FITRES and BONDTWIST around the sugar base bond, the bases are fitted to

the map as well as to each other in a reasonable hydrogen bonding configuration.

The above process introduces some unacceptable stereochemistry, i.e., when the residues are detached and moved and then reconnected, some bond lengths, bond angles and dihedral angles become absurd. Also unusually close contacts between atoms are sometimes introduced. In order to correct these distortions the coordinates are sent to TUCC where the nucleic acid version of the REFINE program is used to idealize the stereochemistry. This is generally done after a session lasting 4-5 hours at the graphic display terminal. These coordinates are then sent back to UNC for reexamination and minor adjustments to ensure a proper fit to the electron density map.

V. REFERENCES

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

Foley, J, and Hogan, R. DIGS- Display Independent Graphics System. Department of Computer Science, University of North Carolina, Chapel Hill, 1975.

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

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