

1977 - 1978  
ANNUAL REPORT

RESOURCE-RELATED RESEARCH  
COMPUTERS AND CHEMISTRY  
Grant No. RR-00612

BIOTECHNOLOGY RESOURCES PROGRAM  
OF THE  
NATIONAL INSTITUTES OF HEALTH

February, 1978

COMPUTER SCIENCE DEPARTMENT  
STANFORD UNIVERSITY

Resource Related Research - Computers and Chemistry  
Stanford University  
NIH/BRP Grant RR-00612

Carl Djerassi, Principal Investigator  
(Social Security No. XXXXXXXXXX)

### Research Highlights (1977-78)

#### 1. Stereochemistry in Structure Elucidation.

The set of computer programs developed at Stanford as tools for molecular structure elucidation have been considerably enhanced by the addition of 3-dimensional structural information. The programs can now deal with some basic geometrical properties of molecules that are essential for understanding their biological significance. Research progress this year has resulted in extensions that allow computation of stereoisomers (alternative structures differing in 3 dimensions but having identical connections among atoms). Thus geometrical variations on structural hypotheses can be presented as well as topological variations.

#### 2. Unified Package for Structure Elucidation.

Significant progress was made in unifying the computer programs for structure elucidation into a coherent package that is easily understood and used by chemists for complex biomolecular structure problems. Powerful tools are now well integrated for defining problem constraints, producing plausible solutions to structure problems, reducing the sets of alternative solutions with information about biosynthetic pathways, testing the alternatives, and suggesting new tests for further discrimination. New tools currently under development will be integrated into this same package.

## Table of Contents

Section		Page
	Subsection	
1.	OVERVIEW OF RESEARCH ACTIVITIES . . . . .	6
2.	STRUCTURE ELUCIDATION PROGRAMS . . . . .	7
	2.1 Stereochemistry in CONGEN . . . . .	7
	2.2 Constraints Interpretation . . . . .	17
	2.3 Experiment Planning Program . . . . .	25
	2.4 The Reaction Chemistry Program . . . . .	35
	2.5 Mass Spectral Prediction and Ranking . . . . .	49
	2.6 Molecular Ion Determination . . . . .	53
	2.7 Congen Improvements . . . . .	60
	2.8 CONGEN Efficiency . . . . .	64
	2.9 CONGEN Reprogramming . . . . .	65
3.	THEORY FORMATION PROGRAMS - Meta-DENDRAL . . . . .	70
	3.1 Incremental Learning . . . . .	70
	3.2 New Capability To Emphasize Discriminatory Power . . . . .	79
	3.3 Improved Ranking Capability . . . . .	80
	3.4 Data Selection Program . . . . .	80
	3.5 Feedback Loops . . . . .	81
	3.6 Program Improvements . . . . .	81
4.	COLLABORATIVE RESEARCH . . . . .	83

4.1	CONGEN Users . . . . .	83
4.2	Marine Natural Products . . . . .	86
5.	Carbon-13 Work . . . . .	92
5.1	Rule Formation Results . . . . .	92
5.2	Adding Stereochemistry to the Rule Language . . . . .	93
5.3	Structure Elucidation . . . . .	94
5.4	Geometric Distortions in Steroids . . . . .	95
6.	DATA COLLECTION AND DATA REDUCTION . . . . .	95
6.1	DENDRAL GC/MS and MS Work . . . . .	95
6.2	Collaborators Receiving the CLEANUP and HISLIB Programs . . . . .	97
7.	APPENDICES . . . . .	102
7.1	Appendix A . . . . .	102
7.2	Appendix B . . . . .	103
	References . . . . .	104