This class of compounds was chosen for two reasons. First, the recent work of Eggert and Djerassi[9a] has yielded a detailed set of predictive rules for the acyclic amines. Secondly, for a given number of Carbon atoms, a saturated, acyclic amine has decidedly more structural possibilities than most other simple types of acyclic organic compounds (for example, stereochemistry aside, there are nearly 15 million C20-amines, but only about 6 million C20-alcohols),[11] and thus the structural analysis of amines represents a particularly challenging problem.

II. DEFINITION OF THE PROBLEM

A fully proton-decoupled, natural-abundance CMR spectrum[8] typically consists of a number of sharp peaks representing the resonance frequencies, in the applied magnetic field, of the various types of Carbon atoms present in the sample. A standard compound, commonly TMS, is usually included in the sample to provide a reference frequency, and the peak positions, or <u>chemical shifts</u>, are measured as fractional deviations from this reference, in parts per million (ppm). Previous investigations have shown that the shift of a particular Carbon is determined by its hybridization and local environment, and thus each shift contains some structural information. There are a few ranges of shifts which are characteristic of certain functional groups, such as C=O or C=C, but aliphatic Carbons in most molecules lie in a broad

spectral region from which detailed structural information cannot be extracted readily.

For acyclic amines,[9a] and a few other types of compounds,[9b-k] there exist predictive rules which allow one to calculate the spectrum of a compound whose structure is known, with a typical accuracy of about 1-2 ppm in a total range of roughly 100 ppm. For these classes, the structure-identification problem could in principle be solved via the generation of all possible structures of a particular type (say, acyclic amines with a particular number of Carbon atoms), the prediction of their spectra, and the comparison of these predictions with the observed spectrum. In fact, Sasaki et al.[6b] have used this procedure in the automated identification of a few small alkanes. For large molecules, though, the number of possible isomers can be overwhelming, and even a very efficient computer program could not carry out such an analysis in a reasonable length of time.

Program AMINE is designed to accomplish the same goal, but in a much more efficient manner. It takes, as its only input data, an observed CMR spectrum, the number of Carbons in the amine, and a goodness-of-fit criterion. The observed spectrum consists of a list of shifts, $\underline{o}=(o_1,\ldots,o_n)$, measured in ppm relative to TMS. Each of these corresponds to one or more Carbons in the sample molecule. Under favorable circumstances,[12] it is possible to determine the number of Carbons corresponding to each observed shift (this will be called the tally of the shift) once the relative peak intensities and the

empirical formula are known. If the tally of a shift is known to be at least 2, 3, etc., then the shift is entered in duplicate, triplicate, etc. in the observed-shift list.   These tallies are not necessary to the program's operation, but even if they are underestimated, they can add considerably to the speed and accuracy of the analysis. The number of Carbons, N_c, in the amine must be greater than, or equal to, the number of shifts in the observed spectrum. Generally, N_c cannot be determined from the CMR spectrum, but must be obtained from some other analytical method such as mass spectroscopy or elemental analysis. The goodness-of-fit criterion, DELTA, which is used in the comparison of $\underline{o}$ to the predicted spectra of molecules or molecular fragments, represents the maximum expected error in the predictive rules. The amine rules are derived, in part, from the alkane rules of Lindeman and Adams,[9d] who note that 95 percent of the studied alkanes have predicted shifts within 1.5 ppm of the observed values. A similar situation exists for the amines, so a value of DELTA = 1.5 ppm has been used in most of this work.

The goal of the program is to find all acyclic, N_c-Carbon amines whose predicted spectra satisfy the following two criteria: a) Every predicted peak must lie within DELTA of one of the observed peaks; and b) Within this limit, the predicted shifts must be assignable to the observed ones in such a way that all of the latter are accounted for.

III. OVERVIEW OF PROGRAM OPERATION

The operation of program AMINE can best be viewed in terms of four interconnected processes; structure generation, pruning, filtering, and spectrum matching. The STRUCTURE GENERATOR builds a pool of increasingly large and complex alkyl chain-ends, and eventually uses these to construct amine molecules. It relies heavily upon the PRUNER to cull from the growing pool any chains which are inconsistent with the observed spectrum, and similarly upon the FILTER to test entire amine molecules. The FILTER also takes care of outputting the acceptable solution structures, and ranking them according to how well they fit the observations. Both the FILTER and the PRUNER use the spectrum MATCHER, which is responsible for the actual comparison of predicted and observed spectra. Each of these processes will be discussed in detail, below.

IV. STRUCTURE GENERATION

The structure generation scheme used in this study, which is related to the enumeration algorithm of Henze and Blair,[13] is applicable only to saturated, acyclic, monofunctional compounds. It is an efficient approach from the standpoint of CMR structural analysis because it rapidly generates substructures which contain a relatively large number of "predictable" Carbons (i. e., those near the ends of alkyl chains), and thus many of these substructures may be ruled out early in the analysis as being inconsistent with the observed data.

At any point in the generation, the STRUCTURE GENERATOR contains a pool of monovalent alkyl radicals which, through pruning (see below) have been found to be consistent with the observed CMR spectrum. The pool initially contains only the $-CH_3$ radical. By attaching one or more of these pool members.(along with an appropriate number of hydrogen atoms) to a central Carbon, it constructs new radicals, each of which is passed to the PRUNER for testing. Any that agree with the observed spectrum are included in the pool, and are subsequently used to construct larger chains. In the final step of the analysis, the STRUCTURE GENERATOR similarly attaches alkyl groups to a central Nitrogen, constructing amine molecules of the proper empirical formula. These it passes to the FILTER for final testing and ranking. At all stages of the generation, tests are made which insure that no radical or amine is considered twice.

As will be discussed below, a given alkyl radical actually undergoes several different tests during pruning, with each test corresponding to a distinct chemical environment in which the chain-end might exist. The STRUCTURE GENERATOR keeps a record of these tests for each pool member, and constantly checks that it is using the radicals in a consistent fashion. If, for example, the PRUNER finds that the ethyl group is consistent with the observed spectrum only if it is attached to Nitrogen in a secondary amine, the STRUCTURE GENERATOR will never construct an n-propyl group, sec-butyl group, or any other radical which contains an ethyl group connected to Carbon. Neither will it generate

primary or tertiary amines with N-ethyl groups.

V. PRUNING

The PRUNER is the real heart of program AMINE. It is responsible for keeping the growing chain-end pool to a manageable size by weeding out alkyl radicals which are inconsistent with the observed spectrum. In testing a particular chain-end, R, shown schematically in Figure 1, the basic question considered by the PRUNER is: "Of all possible sets of CMR shifts which R could produce, is at least one consistent with the observed spectrum?" Actually, the question is somewhat more complex, but this provides a good starting point.

Now, according to the predictive rules,[9a] a Carbon's shift is determined by the structure which surrounds it, up to four bonds away. Further, the effect of a first-row atom which is four bonds removed does not depend upon whether that atom is Carbon or Nitrogen. Thus, because X in Figure 1 must contain at least one such atom (namely Nitrogen), the shifts of C_δ and any Carbons "below" it are completely predictable and independent of the internal structure of X. The shifts of the remaining Carbons, C_γ, C_β and C_α, depend to varying degrees upon the structure of X, with C_α being the most sensitive.

By investigating all possible X structures to a "depth" of four atoms (measured from the R-X bond), the PRUNER could generate an exhaustive list of spectra that R might produce, testing each for

inconsistency with the observed one. Usually, though, X contains enough atoms that there are several hundred of these "depth-4" structures, and the above approach proves to be rather cumbersome. Instead, the PRUNER considers only "depth-3" expansions of X, for which C_β and all Carbons below it are predictable. The shift of C_α is simply ignored, even when a reasonable estimate of its value might be made. This simplification cuts the number of unique X substructures to, at most, 94.

There are two factors which can reduce this number still further. First, some of the substructures may contain too many (or few) Carbons to be consistent with the known atom-count of X. Secondly, there are many cases in which a single predicted spectrum for R may result from two or more related X's. This situation arises because, according to the predictive rules,[9a] the shifts of certain types of Carbons (specifically, those which are four or more bonds from Nitrogen, or three if the degree of the amine is known) are not sensitive to the type or distribution of first-row atoms which are four bonds away, but only to their number. Thus, in the computation of the shift of C_β,

$$-X = -CH\Big\langle {CH\big\langle {{C}\atop{C}} \atop CH_3} $$
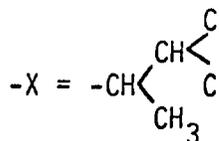
is equivalent to three other structures:

$$-X = -CH\left\langle\begin{array}{c}CH\left\langle\begin{array}{c}C\\N\end{array}\right.\\CH_3\end{array}\right. \quad , \quad -CH\left\langle\begin{array}{c}CH_2-C\\CH_2-C\end{array}\right. \quad , \quad -CH\left\langle\begin{array}{c}CH_2-C\\CH_2-N\end{array}\right.$$
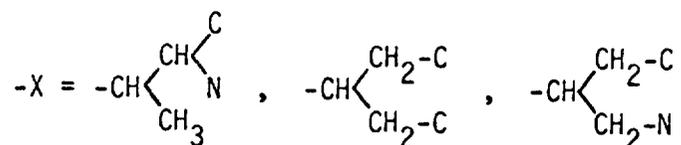
All four may be considered as a single entity, which can be represented as:

$$-X = -CH\left\langle\begin{array}{c}C-\\C-\end{array}\right\}2 \text{ (the no. of non-H atoms)}$$
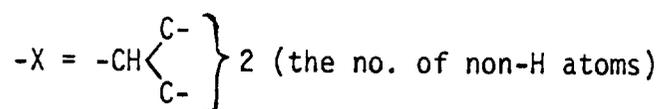
Once such a grouping of X substructures has been done, there remain, at most, 69 cases for the PRUNER to consider. These are summarized in Table 1, where they are further grouped into fifteen classes according to a) the type of atom directly attached to R, b) the degree of that atom and c) if that atom is Carbon, and is attached to Nitrogen, the degree of the amine. The actual purpose of the PRUNER is to consider each of these classes, determining whether at least one class member gives R a predicted spectrum consistent with the observed one, and to return the results of the fifteen class-tests to the STRUCTURE GENERATOR.

The efficiency of this class-by-class investigation can be greatly improved by the inclusion of a hierarchy of pre-tests, each of which is aimed at excluding one or more classes at once. For example, classes

1-12 in Table 1 all have one common feature: The atom to which R is attached is Carbon. Thus, as a pre-test for all twelve classes, the PRUNER treats X as a Carbon whose neighbors are unknown (schematically, X = C-?) and predicts as much of the spectrum of R as possible. If these predictions do not match the observations, it bypasses all further consideration of classes 1-12 and proceeds with the X = N-? pre-test for classes 13-15. Otherwise, it considers a number of more detailed pre-tests, each corresponding to a possible set of neighbors to the central Carbon in X = C-?. The actual hierarchy is outlined in Figure 2. In each of the pre-tests, the local environment of either C_β or C_γ is known to a depth of only three atoms, and hence the corresponding shift cannot be predicted precisely. In most of these cases, the PRUNER can derive upper and lower limits for the shift from the predictive rules. These limits, which define an estimated shift, encompass a relatively small spectral region (0-5 ppm) because the shift of a Carbon is usually not very sensitive to atoms which are four bonds away. Even though the estimated shifts are not exact, they convey useful information to the MATCHER, and thus increase the overall program efficiency.

VI. FILTERING

In the final stages of the analysis, the STRUCTURE GENERATOR constructs amine molecules with N_c Carbons by attaching to a central

Nitrogen, one or more alkyl chains which have survived the pruning process. These amines are passed to the FILTER, which is responsible for calculating their total CMR spectra and, via the MATCHER, comparing these predictions with the observations. If an amine passes the test, the FILTER writes out the structure along with the predicted shifts. It then repeats the spectral comparison using progressively smaller values of DELTA until it finds the smallest value, DELMIN, for which a match still exists. In the event that several solution amines result from a particular run of AMINE, these DELMIN values can be helpful in ranking the candidates according to how well they fit the observed spectrum.

VII. SPECTRUM MATCHING

Eventually, the pruning and filtering processes reduce to problems in spectrum matching. Suppose the MATCHER receives for testing a list of m predicted shifts, some of which may be represented by small spectral regions rather than exact values. Now, the predictive rules are not precise, so each shift is actually associated with a range of acceptable values (given the generic symbol r) whose size is controlled by the input parameter DELTA. This parameter measures the maximum tolerable disagreement between predicted and observed shifts, so the range for a shift, S, extends from S+DELTA to S-DELTA, while that for an estimated shift, bounded above and below by S_u and S_l, extends from S_u+DELTA to S_l-DELTA. It should be noted that, in the latter case,

there are really two factors which contribute to the breadth of the range. One is the basic imprecision in the predictive rules, while the other arises because the PRUNER, in its pre-tests, sometimes calculates shifts for Carbons whose environment is not completely known. The spectrum-matching algorithm, though, makes no distinction between these; It only "knows" that a predicted spectrum consists of a list of ranges. This list will be written as $\underline{r}=(r_1,r_2,\ldots,r_m)$, with u_i and l_i as, respectively, the upper and lower bounds for the range r_i.

The nature of the observed spectrum, $\underline{o}=(o_1,o_2,\ldots,o_n)$, has been discussed in section II. The MATCHER takes these n shifts to be exact, because any estimated uncertainty (usually on the order of 0.1 ppm) in their measurement may be included in the tolerance DELTA. It is the task of the MATCHER to ascertain whether $\underline{r}$ could be a subspectrum of $\underline{o}$, or if $m=N_c$ (N_c being the number of Carbons in the amine), whether $\underline{r}$ could be interpreted as $\underline{o}$.

If, for a range r_i, there exists an observed shift o_j such that $u_i \geq o_j \geq l_i$, then it will be said that r_i can be assigned to o_j. The simplest test of agreement between $\underline{r}$ and $\underline{o}$ involves checking that each r_i in $\underline{r}$ can be assigned to at least one o_j in $\underline{o}$. This test does not consider the important condition that, eventually, all shifts in $\underline{o}$ must be used, and therefore a stronger test can be defined.

If every Carbon in the molecule gives a different observed shift, or if an analysis of peak intensity data gives the tally of each peak, then $n=N_c$. In this case, it is clear that no two predicted shifts can

be assigned to the same o_j. Thus, referring to Figure 3, $\underline{r}$ does not match $\underline{o}$ even though the simple test is not violated, because r_1 and r_3 must both be assigned to o_2. In more complicated cases, each of several r_i's may be assignable to two or more o_j's, and vice versa, so the application of this test in an efficient manner can present difficulties. Fortunately, simple matching theory,[14a] an outgrowth of the mathematical field of graph theory, provides a general method (see MATCHING ALGORITHM, below) of finding the maximum number, M, of ranges which can be assigned to the elements of $\underline{o}$ without duplication. Clearly $\underline{r}$ cannot match $\underline{o}$ if this number is less than m.

There may be cases, though, in which complete tally information is unavailable, which means that the number of observed shifts, n, is smaller than the number of Carbons, N_c. In such cases, there are (N_c-n) "extra" shifts which lie somewhere beneath the n observed ones, but there is no way of determining where they belong. It is still possible to strengthen the simple test, but here, the additional constraint is that the predicted spectrum, once assigned, can have no more than (N_c-n) "extra" peaks, either. If the simple test is passed, then every r_i can be assigned to at least one o_j. However, M is the maximum number of ranges which can be assigned to $\underline{o}$ without duplication, so (m-M) must be the number of "extra" ranges in $\underline{r}$. The condition that strengthens the simple test here, then, is $(m-M) \leq (N_c-n)$. Because M cannot exceed m, this condition reduces to the previous one (m=M) when $n=N_c$.

The above spectrum-matching scheme is useful not only in the current study, but for general cases in which a set of predicted CMR shifts of variable uncertainty is to be compared with an observed spectrum with, perhaps, incomplete intensity information.

VIII. MATCHING ALGORITHM

The algorithm for determining M is related to the so-called Hungarian method[14b] of simple matching, but takes advantage of certain special features of the spectrum-matching problem. It may be described briefly as follows: Begin with M=0 and process the o_j's in algebraic order, beginning with the largest. For each o_j, scan the ranges r_i looking for those which satisfy $u_i \geq o_j \geq l_i$, but which have not yet been assigned. If there are none, proceed to the next o_j. If there is just one, assign it to o_j, increment M by 1 and proceed to the next o_j. If there are several, assign the one with the largest lower limit (l_i) to o_j, increment M by 1 and proceed to the next o_j.

It is possible to prove that this gives the maximum matching between $\underline{r}$ and $\underline{o}$, but a presentation of our proof is beyond the scope of this paper.

IX. RESULTS

Program AMINE has been implemented on the IBM 360/67 and DEC PDP-10

computers. Any mention of timing in the following discussion refers to total execution time (central processor time plus "wait time") on the former machine.[15] The program requires about 35 thousand words of storage. A sample of the program's output is shown in Figure 4.

The only large set of amine CMR spectra available in the literature is that given by Eggert and Djerassi,[9a] who used it in the derivation of predictive rules. The set consists of 102 amine spectra, including both shifts and tallies. Three of these spectra correspond to diastereomeric mixtures, and these are not suitable for testing AMINE, because the program assumes that the input spectrum corresponds to a pure compound. Neither is tridodecylamine because it exceeds the maximum number of Carbons (currently 24) allowed by the program. The remaining 98 amines were used in the testing of the program.

Some experimentation indicated that DELTA=1.5 ppm was small enough for efficient and selective program operation, yet large enough that about 95% of the test cases gave the correct solution among the output structures. Increasing DELTA by 50% to 2.25 ppm slowed the program by a factor of 2-4, but AMINE always obtained the correct structure with this higher DELTA value. Generally, shift tallies were found to be unnecessary for amines containing fifteen or fewer Carbons, but for larger molecules, the analyses proved to be excessively costly unless all of the Carbons were identified in the observed spectrum.

With DELTA=1.5 ppm, and using tallies for the amines with sixteen or more Carbons, the program obtained only one answer, the correct one,

for the 88 amines listed in Table 2. The six cases summarized in Table 3 gave from two to seven solutions, with the correct structure ranked (see section VI.) first or tied[16] for first. For three of these six, the inclusion of tallies ruled out the incorrect answers. Four amines gave no solutions with DELTA=1.5 ppm. These were rerun using DELTA=2.25 ppm, and tallies were included to offset the longer running-times. As indicated in Table 4, three of these runs gave only the correct answer, while the fourth yielded two equally ranked solutions, including the correct one.

These analyses required from 0.02 to 100 seconds of computer time, with a typical 10- or 11-Carbon amine using about 1-2 seconds. In none of the runs was an incorrect solution obtained without the accompanying correct one, and in only four cases was it necessary to use the larger DELTA value. The results for the eight amines containing sixteen or more Carbons are especially encouraging: In a reasonable length of time, the program was able to select the correct structure, along with very few others, from an "isomer space" containing from about 300,000 (for N_c=16) to about 700,000,000 (for N_c=24) members.[11]

The above results are biased to some extent because the amines used for testing the program are the same ones used by Eggert and Djerassi in the predictive-rule formation. As a test of the generality of the program, analyses have been run on the spectra of four "unknown" amines which do not appear in the original list. The results of these tests cases are summarized in Table 5. The spectra of the two 13-Carbon

amines were analyzed using DELTA=1.5 ppm, and no attempt was made to include tallies. Only the correct structure was obtained in these cases. For the two 20-Carbon amines, tallies were measured under special experimental conditions (see below). With DELTA=1.5 ppm, one of these gave two equally ranked structures, while the other gave none. A rerun of the second case with DELTA=2.25 ppm yielded five solutions with the correct one ranked as tied for second. This is the only case in which the ranking procedure favored an incorrect answer over the correct one, but here, as in most of the other multiple-result runs, the incorrect structures are sufficiently different from the correct one that they should be distinguishable by mass-spectroscopic techniques.

X. CONCLUDING COMMENTS

Two major conclusions result from this study. First, the CMR spectrum of an acyclic amine appears to be highly characteristic of the structure of the amine. For example, only one of the nearly 15 million[11] structural isomers of $C_{20}H_{43}N$ gives a predicted spectrum which matches the observed spectrum of N-butyldi(2-ethylhexyl)amine. Thus it can be concluded that CMR data do indeed contain a tremendous amount of structural information. Secondly, it has been found that efficient methods for extracting this information exist, and can be implemented on the digital computer.

There is no reason to believe that these conclusions are peculiar

to the acyclic amines. The computational techniques outlined in this paper can readily be generalized to other classes of saturated, acyclic, monofunctional compounds: To do so for a particular class, one needs to obtain an accurate set of predictive rules, and, perhaps, to modify the pruning process slightly to account for special features of those rules. Such rules already exist for alkanes[9c,d], alkenes[9e], and alcohols[9b], and as research in CMR spectroscopy progresses, further sets should become available. Extensions to polyfunctional and/or cyclic classes would also require more sophisticated structure-generation methods, but these are available.[3,6a,17]

In short, it appears that the computerized analysis of CMR spectra holds great promise as an accurate and selective tool in the identification of unknown compounds.

XI. EXPERIMENTAL

The four "unknown" amines were prepared, and their proton noise-decoupled CMR spectra obtained, using previously described techniques.[9a] The spectra of the two 20-Carbon amines were also run in the presence of chromium acetylacetonate, and the integrated intensities from these were used to determine the peak tallies.[18] The observed shifts for the four amines are given below, in ppm downfield of internal TMS. The estimated uncertainty in these shifts is 0.1 ppm. For the two 20-Carbon amines, tallies are included in parentheses.

N-(3-methylbutyl)-2-ethylhexylamine;53.6, 48.6, 39.8, 39.6, 31.7, 29.3, 26.3, 24.8, 23.2, 22.8, 14.1, 11.0.

N-(3-methylbutyl)-1,5-dimethylhexylamine;53.5, 45.6, 39.9, 39.4, 37.8, 28.1, 26.4, 24.0, 22.7, 20.6.

N-(2-ethylhexyl)-N-(3-methylbutyl)heptylamine;59.6(1), 55.0(1), 53.1(1), 38.1(1), 36.8(1), 32.3(1), 31.8(1), 29.7(1), 29.4(1), 27.9(2), 26.5(1), 24.9(1), 23.6(2), 23.0(2), 14.3(2), 11.0(1)

N-pentyl-N-(3,3-dimethylbutyl)-3,5,5-trimethylhexylamine; 54.5(1), 52.5(1), 51.9(1), 50.1(1), 40.9(1), 31.1(1), 30.2(3), 30.0(1), 29.8(1), 29.7(3), 27.7(2), 22.9(2), 14.1(1).

XII. ACKNOWLEDGEMENTS

acknowledged.

References

1. For part X, see D. H. Smith, B. G. Buchanan, W. C. White, E. A. Feigenbaum, J. Lederberg, and C. Djerassi, submitted for publication in Tetrahedron.

2. National Institutes of Health postdoctoral Fellow, 1972-1973.

3. (a) B. G. Buchanan, A. M. Duffield, and A. V. Robertson in "Mass Spectroscopy: Techniques and Applications," ed. G. W. A. Milne, Wiley and Sons, New York, 1971, p. 121; (b) D. H. Smith, B. G. Buchanan, R. S. Engelmore, A. M. Duffield, A. Yeo, E. A. Feigenbaum, J. Lederberg, and C. Djerassi, J. Amer. Chem. Soc., 1972, 94, 5962, and previous papers in the series.

4. A. Buchs, A. M. Duffield, G. Schroll, C. Djerassi, A. B. Delfino, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum, and J. Lederberg, J. Amer. Chem. Soc., 1970, 92, 6831.

5. (a) H. S. Herze, R. A. Hites, and K. B. Biemann, Anal. Chem., 1971, 43, 681; (b) L. R. Crawford and D. J. Morrison, ibid., 1971, 43, 1790; (c) D. H. Smith, ibid., 1972, 44, 536; (d) P. C. Jurs, ibid., 1971, 43, 1812, and references cited therein.

6. (a) S.-I. Sasaki, Y. Kudo, S. Ochiai, and H. Abe, Mikrochimica Acta [Wien], 1971, 726; (b) S.-I. Sasaki, S. Ochiai, Y. Hirota, and Y. Kudo, Japan Analyst, 1972, 21, 916.

7. H. Abe and S.-I. Sasaki, The Science Reports of the Tohoku University, Series I, 1972, 55, 63.

8. For a general discussion of CMR spectroscopy, see "Carbon-13 Nuclear Magnetic Resonance for Organic Chemists," G. C. Levy and G. L. Nelson, Wiley-Interscience, New York, 1972.

9. (a) H. Eggert and C. Djerassi, J. Amer. Chem. Soc., in press; (b) J. D. Roberts, F. J. Weigert, J. I. Kroschwitz, and H. J. Reich, ibid., 1970, 92, 1338; (c) D. M. Grant and E. G. Paul, ibid., 1964, 86, 2984; (d) L. P. Lindeman and J. Q. Adams, Anal. Chem., 1971, 43, 1245; (e) D. E. Dorman, M. Jautelat, and J. D. Roberts, J. Org. Chem., 1971, 36, 2757;

-23-

(f) D. K. Dalling and D. M. Grant, J. Amer. Chem. Soc., 1967, 89, 6612;
(g) M. Cristl, H. J. Reich, and J. D. Roberts, ibid., 1971, 93, 3463;
(h) D. E. Dorman, S. J. Angyal, and J. D. Roberts, ibid., 1970, 92, 1351;
(i) J. K. Crandall and S. A. Sojka, ibid., 1972, 94, 5084;
(j) W. R. Woolfenden and D. M. Grant, ibid., 1966, 88, 1496;
(k) F. J. Weigert and J. D. Roberts, ibid., 1970, 92, 1347, and references cited therein.

10. Copies of the program, along with sample input decks, may be obtained from the authors.

11. These isomer counts were computed using the enumeration algorithm of Henze and Blair, Reference 13.

12. G. N. La Mar, J. Amer. Chem. Soc., 1971, 93, 1040.

13. H. R. Henze and C. M. Blair, J. Amer. Chem. Soc., 1931, 53, 3042 and 3077.

14. (a) "The Theory of Graphs and its Applications," C. Berge, Wiley and Sons, New York, 1964, p. 92;
(b) ibid., p. 99.

15. On the PDP-10, the program runs more slowly by a factor of about four (central processor time only).

16. Two structures are considered to be tied when their DELMIN values differ by 0.1 ppm or less.

17. Exhaustive, irredundant methods for the generation of cyclic structures have recently been developed by L. M. Masinter and N. S. Sridharan as part of our Heuristic DENDRAL project. A manuscript describing their work is in preparation.

18. S. Barcza and N. Engstrom, J. Amer. Chem. Soc., 1972, 94, 1762.

Table 1. The Substructures X Considered by the PRUNER.

Class	"Depth 3" X structure(s)		Class	"Depth 3" X structure(s)

Class 1: $R-CH_2-C-$ } 1,2,3

Class 2: $R-CH_2-NH_2$

Class 3: $R-CH_2-NH-C$

Class 4: $R-CH_2-N\Big\langle{{C}\atop{C}}$

Class 5: $R-CH\Big\langle{{C-}\atop{C-}}$ } 1,2,...,6

Class 6: $R-CH\Big\langle{{NH_2}\atop{C-}}$ } 0,1,2,3

Class 7: $R-CH\Big\langle{{NH-C}\atop{C-}}$ } 0,1,2,3

Class 8: $R-CH\Big\langle{{N{\langle}^{C}_{C}}\atop{C-}}$ } 0,1,2,3

Class 9: $R-C\Big\langle{{C-}\atop{C-}\atop{C-}}$ } 1,2,...,9

Class 10: $R-C\Big\langle{{{C-}^{NH_2}}\atop{C-}}$ } 0,1,...,6

Class 11: $R-C\Big\langle{{{C-}^{NH-C}}\atop{C-}}$ } 0,1,...,6

Class 12: $R-C\Big\langle{{{C-}^{N{\langle}^{C}_{C}}}\atop{C-}}$ } 0,1,...,6

Class 13: $R-NH_2$

Class 14: $R-NH-C-$ } 0,1,2,3

Class 15:

$R-N\Big\langle{{C-}\atop{C-}}$ } 0,1,5,6

$R-N\Big\langle{{CH{\langle}^{C}_{C}}\atop{CH_3}}$

$R-N\Big\langle{{CH_2-C}\atop{CH_2-C}}$

$R-N\Big\langle{{C{\langle}^{C}_{C}}\atop{CH_3}}$

$R-N\Big\langle{{CH{\langle}^{C}_{C}}\atop{CH_2-C}}$

$R-N\Big\langle{{C{\langle}^{C}_{C}}\atop{CH_2-C}}$

$R-N\Big\langle{{CH{\langle}^{C}_{C}}\atop{CH{\langle}^{C}_{C}}}$

P-72

Table 2. Cases for which AMINE obtained only the correct structure
          using DELTA = 1.5 ppm and, except as noted, no tallies.

Amine (prefix only)

methyl
ethyl
propyl
isopropyl
trimethyl
butyl
sec-butyl
isobutyl
tert-butyl
diethyl
pentyl
1-methylbutyl
2-methylbutyl
3-methylbutyl
2,2-dimethylpropyl
N-methyl-sec-butyl
N-methyl-tert-butyl
N-methyldiethyl
hexyl
1,3-dimethylbutyl
1,2,2-trimethylpropyl
2,2-dimethylbutyl
dipropyl
diisopropyl
N-ethylbutyl
N-ethyl-sec-butyl
triethyl
N,N-dimethyl-sec-butyl
N,N-dimethyl-tert-butyl
heptyl
1-methylhexyl
1-ethylpentyl
1,3-dimethylpentyl
N-methylhexyl
N-isopropylbutyl
N-isopropyl-sec-butyl
octyl
1-methylheptyl
2-ethylhexyl
1,5-dimethylhexyl
1,1,3,3-tetramethylbutyl
dibutyl
diisobutyl
N-ethylhexyl
N,N-dimethylhexyl
N,N-diethylbutyl
N,N-diethyl-sec-butyl

Amine (prefix only)

N-ethyldiisopropyl
nonyl
N-propylhexyl
N-sec-butylpentyl
N-sec-butyl-3-methylbutyl
N-tert-butyl-3-methylbutyl
N-methyl-1,1,3,3-tetramethyl-
     butyl
tripropyl
decyl
dipentyl
N-butylhexyl
N-tert-butylhexyl
N-sec-butyl-3,3-dimethylbutyl
di(3-methylbutyl)
N-ethyldibutyl
N-ethyldibutyl
N,N-diisopropylbutyl
N-pentylhexyl
N-butyl-1-methylhexyl
N-pentyl-1,3-dimethylbutyl
N-(3,3-dimethylbutyl)pentyl
N-butyl-1-ethylpentyl
N-methyl-N-butylhexyl
N-propyldibutyl
N-isopropyldibutyl
N-(1,3-dimethylbutyl)hexyl
tributyl
N-ethyldipentyl
N-tert-butyldibutyl
N,N-dibutyl-3-methylbutyl
N,N-dibutylhexyl
N,N-dibutyl-3,3-dimethylbutyl
N-sec-butyldipentyl
N,N-dipentyl-1-methylpentyl
tripentyl
tri(3-methylbutyl)

Using tallies:

di(2-ethylhexyl)
N,N-dipentyl-1,3-dimethylbutyl
N,N-dibutyl-1,1,3,3-tetramethyl-
     butyl
trihexyl
N-butyldi(2-ethylhexyl)

Table 3. Cases for which AMINE obtained two or more structures
using DELTA = 1.5 ppm and, except as noted, no tallies.

Amine (prefix only)	Solutions (prefix only)	Rank
dihexyl	dihexyl	tied
	N-pentylheptyl [a]	tied
N-pentyl-1,1,3,3-tetra-methylbutyl	N-pentyl-1,1,3,3-tetra-methylbutyl	1
	N-tert-butyl-1,1-dimethyl-heptyl[a]	2
N-(1-ethylpentyl)-1-propyl-butyl	N-(1-ethylpentyl)-1-propyl-butyl	tied
	N-(1-ethylbutyl)-1-propyl-pentyl	tied
N,N-dibutylheptyl	N,N-dibutylheptyl	1
	N-butyl-N-pentylhexyl[a]	2
dioctyl[b]	dioctyl	tied
	N-heptylnonyl	tied
	N-hexyldecyl	tied
trioctyl[b]	trioctyl	tied
	N-heptyl-N-octylnonyl	tied
	N,N-diheptyldecyl	tied
	N-hexyldinonly	tied
	N-hexyl-N-octyldecyl	tied
	N-hexyl-N-heptylundecyl	tied
	N,N-dihexyldodecyl	tied

a) The use of tallies excludes these structures.
b) Tallies were used in these runs.

Table 4. Cases for which AMINE found no structures using
DELTA = 1.5 ppm. The correct solutions appeared
when DELTA was increased to 2.25 ppm and tallies
were included.

Amine

1-isopropylhexylamine

N-pentyl-1,2,2-trimethylpropylamine[a]

N-butyl-N-(1,2,2-trimethylpropyl)pentylamine

N-butyl-N-pentyl(1,1,3,3-tetramethylbutyl)amine

a) A second structure, equally ranked, was found in this
   case: N-propyl-N-(1,2,2-trimethypropyl)hexylamine.