C13Shift: A Computer Program for the Prediction of ^{13}C NMR Spectra Based on an Open Set of Additivity Rules

Ernö Pretsch,* Andràs Fürst, Martin Badertscher, and Renate Bürgin

Department of Organic Chemistry, Swiss Federal Institute of Technology, CH-8092 Zurich, Switzerland

Morton E. Munk

Department of Organic Chemistry, Arizona State University, Tempe, Arizona 85287-1604

Received February 10, 1992

The strategies to increase the range of applicability of a recently described program C13Shift are discussed. The modularity of the approach and the independence of the parameters allow the addition of new rules and new parameters. The program has been interfaced to a molecular assembler so that generated structures can be ranked according to the deviations between observed and predicted chemical shifts.

INTRODUCTION

Since the early years of the use of NMR spectroscopy for structure elucidation, additivity rules have been proposed to estimate the ^1H NMR and ^{13}C NMR chemical shifts for a large number of substance classes.[1] Although the models applied vary considerably, they have some common properties. First, most of the models proposed are linear. Nonlinearities have been taken into account by adding some "correction terms" in several cases. Second, substituents were treated as a whole in most of the instances, i.e., increments were group increments. Often increments were determined simply from the chemical shifts of the corresponding unsubstituted and monosubstituted skeleton. The most important exception is the Grant/Paul additivity rule for the estimation of ^{13}C NMR chemical shifts in aliphatic hydrocarbons.[2] Here increments are used for the individual carbon *atoms* up to a distance of five bonds apart from the center of interest. The extensions[1,3] of this rule for substituted aliphatic compounds use partly atomic increments, e.g., for O in ethers and alcohols. Group increments were, however, introduced for functional groups which cannot be described adequately with atomic increments (e.g., for O—C=O). The groups are a kind of superatoms. When counting bonds between atoms, the number of bonds within the group are added to the ones encountered otherwise. Both atomic and group increments have an easily interpretable meaning for the chemist.

Several computer programs were described to estimate the chemical shift of sp^3-hybridized carbon atoms on the basis of additivity rules.[4-10] The first general solution covering a wide variety of carbon atom types was published recently.[11] This program automatically selects and applies the estimation model for each carbon atom of the molecule entered. It treats both atomic and group increments for sp^3-hybridized carbon atoms and group increments for all other carbon atoms.[12,13]

Another approach based on linear models employs soft modeling to select the types of model parameters. In the first modeling step a larger number of possible descriptors are derived from topological as well as three-dimensional representations of the structure. The most relevant parameters are then selected by multiple linear regression analysis.[14-18] These models are usually tailored for a relatively narrow class of compounds like alkylcyclopentanes or cyclic alcohols on the basis of up to several hundred compounds with assigned chemical shifts. They allow the most accurate predictions currently available: standard errors of 0.5–2 ppm were reported—which is sometimes even lower than variations[19] due to changes of experimental conditions! The first step of the estimation is a molecular mechanics relaxation of the structure since a three-dimensional structure is needed for the prediction. Parameter sets became available for nearly 100 compound classes. An automatic selection of the model from a database is possible.[18]

The third basic approach uses assigned spectra of a database for prediction purposes.[20-23] Linear codes of atom-centered fragments of every carbon atom in the whole database are generated and collected in an ordered list. For the estimation, the same kind of atom-centered fragments of the proposed structure are generated. The chemical shifts of the corresponding entries in the database serve as estimates. If the corresponding atom-centered fragment is not available in the database, the program can "interpolate" by using the atom-centered fragments which would be next to the unknown center in the list used for the prediction. A software package is available on a host[24] with numerous users worldwide.

THE C13SHIFT PROGRAM

The C13Shift program automatically selects and applies the additivity rules for the individual carbon atoms of the structure entered.[12] The connectivity matrix used for the internal representation of the molecule can be generated on the basis of a linear code, similar to the SMILES code.[25] Alternatively, the output of a molecular assembler program[26] can be directly used as structure input.

A hierarchical list contains all substructures for which additivity rules are available. The atoms of the compound entered are assigned to one of the substructures. Every atom can only be assigned to one substructure, so that after a successful assignment of a set of atoms they are no longer considered for substructure assignment. For each set of atoms belonging to one substructure, all other atoms of the molecule are considered as substituents of this substructure. Values for their increments are stored in the corresponding records of the substituent file. Group increments are used for all sp^2- and sp-hybridized carbon atoms, and both atomic and group increments are applied for all sp^3-hybridized carbon atoms.

As an example, the chemical shift estimation of atropine is shown in Figure 1. The linear code input is displayed on the top of the window. Below it the constitution is depicted

0095-2338/92/1632-0291$03.00/0

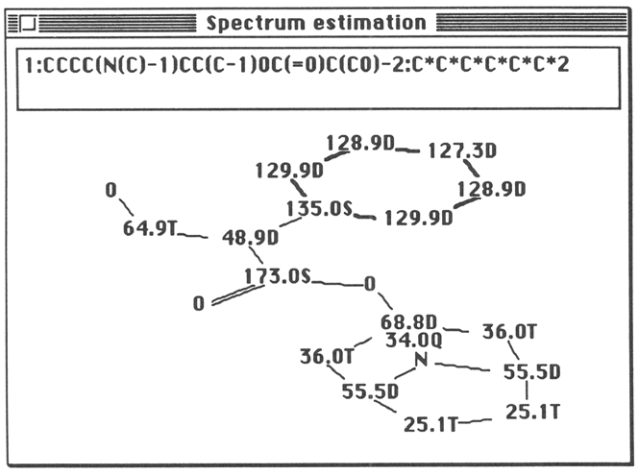

Figure 1. Estimation of the ^{13}C NMR spectrum of atropine with the program C13Shift.

```
C13 spectrum of 1:N*C(I)*C(OCC)*C*C*C*1

Substructure assignment trials:
     2 for structure pyridine
   base value for atom# 1 is 149.8 from pyridine
   base value for atom# 3 is 123.6 from pyridine
   base value for atom# 7 is 135.7 from pyridine
   base value for atom# 8 is 123.6 from pyridine
   base value for atom# 9 is 149.8 from pyridine
Atom #0 in substructure pyridine
Atom #1 in substructure pyridine
substituent file pyridine.a with 29 records read in
Corresponding substructure:benzene
Corresponding atoms: 0->0. 1->1. 3->2. 7->3. 8->4. 9->5.
substituent file benzene.a with 131 records read in
   substituted by   I incr: -31.2(1) 8.9(3) 1.6(7) -1.1(8) 1.6(9)
Atom #3 in substructure pyridine
substituent file pyridine.b with 27 records read in
   substituted by   O-C incr: -12.5(1) 31.6(3) -15.7(7) 0.2(8) -8.4(9)
Atom #7 in substructure pyridine
Atom #8 in substructure pyridine
Atom #9 in substructure pyridine
Atom #5 aliphatic order=2
substituent file sp3shift.a with 81 records read in
  alpha substitution by O increment: 49.0 ppm
    steric corr order=2 i=2 corr= 0.0
  alpha substitution by C increment:  9.1 ppm
    steric corr order=2 i=1 corr= 0.0
  beta substitution by 1:C*A*A*A*A*A*1 increment:  9.3 ppm
  delta substitution by I increment: -0.9 ppm
Atom #6 aliphatic order=1
  alpha substitution by C increment:  9.1 ppm
    steric corr order=1 i=2 corr= 0.0
  beta substitution by O increment: 10.1 ppm
```

Figure 2. Part of the protocol of the estimation of 2-iodo-3-ethoxypyridine.

together with the estimated shift values and with symbols of the multiplicity in the off-resonance decoupled spectrum. Alternatively a table output can be presented. A detailed protocol (not shown in Figure 1) is automatically generated by the program during the estimation procedure.

One goal of the present program was to achieve a wide range of applicability. Besides a substantial increase of the available parameter sets, several strategies were implemented to further widen the scope of the C13Shift program. These strategies include the substitution of a missing parameter by the one of a similar substituent. Embedded substituents, i.e., substitutents being part of another, are considered as similar substituents. In case no embedded substituents can be found at all, values for the same (or embedded) substituent of a similar substructure are used. Thus, for example, if 2-iodo-3-ethoxypyridine were the structure entered, the program would apply the increment of a *methoxy* group in 3-position of pyridine, since no ethoxy is available (embedded substituent), and that of a iodine substituent of *benzene*, since no value for pyridine is available (see Figure 2). Both of these strategies would be applied by an experienced chemist manually. Whereas the replacement of increments for missing

```
indole
1:C-C=C-N-C(*1)*C*C*C*C*1
  c b a 1 h     g f e d
  * * * * *
  *       *     * * * *
1237 1018 1270 1199 1190 1211 1104 1348
```

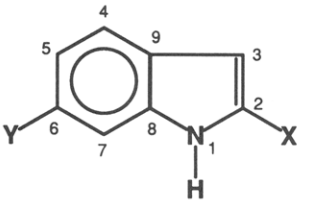

Figure 3. Record of the substructure indole in the substructure file. The lines 4 and 5 define pyrrole (for substituents 1-3, e.g., 2-X) and benzene (for substituents 4-7, e.g., 6-Y) as corresponding substituents.

substituents by increments of embedded groups is a fully automatic procedure hard-coded in the program, the replacement of skeletons by similar skeletons is defined in the substituent files and is thus under user control (see below).

Using some 3000 increments[12,13] and the strategies described to increase the range of applicability, 97% of over 160 000 chemical shifts in a database[22] (extended version with over 25 000 spectra) could be estimated. The mean deviation between predicted and observed values was –0.29 ppm with a standard deviation of 5.5 ppm. The real standard deviation is probably somewhat lower since the automatic evaluation was biased by the fact that no stereochemical information is available in the database. The program distinguishes between cis and trans isomers of double bonds as well as axial and equatorial substitution of six-membered rings.[11] In addition, erroneous assignments in the database would typically increase the deviation between observed and estimated values.

C13Shift has now been in routine use for 3 years by chemists and students in our laboratory. It proved to be a valuable tool. Some of the capabilities, however, have gained too little recognition even by routine users of the program. These will be discussed in the next sections of this paper.

PARAMETER SET USED BY THE C13SHIFT PROGRAM

No parameters are hard-coded in the program text but collected in external files. Thus parametrization is fully under the user's control. The user can modify or extend the parameter sets or even introduce new additivity rules for skeletons not considered in the present version, as long as they fit into the general scheme of the linear model applied. Chemists working in specialized fields have much more information about the compounds of their interest than was incorporated into the general purpose parameter set. The users are therefore strongly encouraged to formulate this information as an extension to the current database. Some of the parameters were adjusted on the basis of a large number of observations. Their change on the basis of some new observations is, therefore, questionable. There are three parameter files: a substructure file, a substituent file, and a correction file. The *substructure file* contains the name, constitution, and symmetry properties of a substructure together with the chemical shifts. Corresponding substructures can be defined as an option (cf. Figure 3). If no substituent increments are available for a substructure, those of a corresponding substructure are used together with base values of the substructure. Corresponding substructures may have the same number of atoms as the substructure (e.g., benzene as a corresponding substructure for pyridine) or less atoms as

```
pyrimidine
1:N*C*N*C*C*C*1
  1 a 1 b c b
  ! * C ! * *
  C ! ! * ! !
1580 1564 1214
```

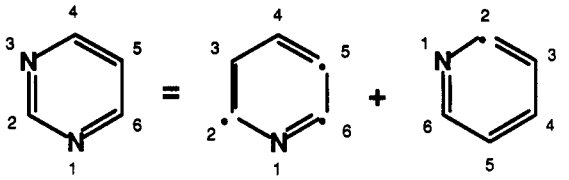

Figure 4. Overlay of a corresponding substructure (pyridine) several times on the parent structure (pyrimidine). Only the atoms of pyridine marked with a dot are used for the estimation. The syntax is shown on the top.

the substructure, i.e., they may be sub-substructures. Thus, pyrrole and benzene can be corresponding substructures for indole (Figure 3). For a 2,5-disubstituted indole, the increments of the 2-substituent X for pyrrole (fourth line in Figure 3) will be used to estimate the shifts of the atoms 1-3, 8, and 9 and the increments for benzene will replace the increments of the 6-substituent Y for the atoms 6–9 if these substituents are missing in the indole substituent file. The contributions of X to the shifts 4–7 and those of Y to 2–3 will be neglected.

A corresponding substructure may consist of the same atoms as the parent structure or some of the atoms may be changed (typically to C or N, e.g., when benzene is used as a corresponding substructure of pyridine). In some cases, to perform the best estimation and/or to keep the symmetry of the parent substructure, a corresponding substructure may be overlaid to the same parent structure several times. In this case it might be necessary to forbid the estimation of a carbon shift by one of the overlays. If, for example, the chemical shifts of a substituted pyrimidine are estimated, the increments of pyridine can be used (cf. Figure 4). The substituents of the 2,4 and 6 positions (C next to N) are approximated by the effects of the 2-substituents in pyridine. If pyridine is first overlaid on pyrimidine, only its substituents on the 2,5 and 6 positions are therefore considered but not those on the 4 position. In the second overlay (line 5 in Figure 4) the remaining 4 position of pyrimidine will be taken into account. The syntax allows the proper treatment in such cases: the marker sign "!" (cf. Figure 4) has the desired effect.

The chemical shifts of sp^3-hybridized carbon atoms are estimated with atomic or superatomic increments added to the base value of –2.3 ppm, the chemical shift of methane. For many skeletons this approach would lead to a systematic error which mainly depends on the kind of skeleton but scarcely on the kind of substitution. Therefore so-called virtual substructures are implemented in the model. For examples see Scheme 1 in ref 12. They are called virtual, since no parameter records belong to them in the substituent file. For sp^3-hybridized carbon atoms of virtual substructures the same increments are used as for all other sp^3-hybridized carbon atoms. For the estimation of a chemical shift of a carbon atom within the virtual substructure *all* other atoms of the molecule, including those of the substructure itself, are considered as substituents. The base-shift values (stored with the virtual substructure) are chosen to reproduce the exact chemical shifts of the unsubstituted skeleton. Non-sp^3-hybridized atoms of virtual substructures can be assigned to any other corresponding substructures in the usual manner.

The *substituent file* contains increments for substituents on every position of a skeleton (as defined in the skeleton file). Increments are given for every carbon atom of the skeleton for each substituent position. Because of the principle of substituting missing substituents by embedded ones, the substituents have to be listed in a hierarchical order. For example, the alkyl substituents are in the order *tert*-butyl, isopropyl, butyl, propyl, ethyl, and methyl. Thus, an isobutyl group (for which no increment is given in this example) will be replaced by isopropyl and not by propyl.

The *correction file* contains terms accounting for nonlinearities in case of sp^3-hybridized carbon atoms. One of them is the so-called α-corrections.[12] An additive term is assigned to a number of simultaneous occurrences of α-substituents. Only one term is applied, the one which is first encountered in the hierarchically ordered list. Another set of terms takes into account the influence of steric relation of γ-substituents.[1] The term will be selected automatically in case of double bonds (cis/trans) and six-membered rings (axial vs equatorial substituents). The linear model is a reasonable approach for β-substituents, so no correction terms are needed.

The nonlinearities induced by various branching of a carbon center were originally taken into account by so-called steric corrections[2] for aliphatic hydrocarbons. These terms are later extended to other atoms than carbon.[1,3] In the C13Shift program individual sets of steric corrections are implemented for three- to six-membered rings and for acyclic and larger ring systems.

PARAMETER EDITING BY THE USER

Adding New Substituent Increments for Existing Skeletons. It is straightforward to insert additional increments into the existing substituent files. The only precaution to be kept in mind is the hierarchical order. If, for example, increments are listed for propyloxy and methoxy, new increments for ethoxy have to be inserted between the two, otherwise some of the increments will never be used.

Modifying Existing Substituent Increments. Some of the existing increments were adjusted on the basis of a large number of observations (cf. refs 12 and 13). Their change on the basis of a few new data is not advisable. New data can, however, be used to modify existing increments which were derived only on the basis of a few observations. The number of occurrences of the individual substituents are given in the original papers.[12,13]

Correction Terms. If new data are available to introduce missing α-correction terms, this can safely be done. Again, the hierarchical order has to be kept in mind. Existing values should only be modified if the new values are based on a sufficiently large number of observations.

The steric correction terms for *ring systems* are only rough estimates since sufficient data were not available for a reliable estimation. Those of the noncyclic systems were carefully investigated on the basis of a large data set and should therefore not be changed.

Adding New Virtual Substructures. There is no limitation to adding new virtual substructures, i.e., substructures for which at least partly the sp^3-hybridized carbon atom increments have to be used with modified base values. It is necessary to use virtual substructures for reliable chemical shift estimations in strained condensed ring systems and in five-membered rings. Again the hierarchical order of the substructures has to be kept in mind.

The general hierarchy is to proceed from larger to smaller substructures. There is, however, one exception from this rule in the case of virtual substructures: atoms will be reassigned to another virtual substructure if they belong to a

smaller ring than the one in the originally assigned substructure (for further explanation see ref. 12). The reason for this is that the magnitude of the necessary correction terms increases with decreasing ring size. It is also possible to use substituted skeletons as virtual substructures. This has been done in the case of cyclopropane substituted with double or triple bonded substituents like C=O or C≡N. This became necessary because the usual linear model led to consistently large errors for the atoms of the three-membered ring which are not bonded to the substituent.

Adding New Skeletons with Substituent Parameters. There is no limitation to adding new skeletons with their parameter sets. Again the only precaution is that the hierarchical order has to be considered both for the position of the skeleton in the substructure file and for the substituents within the substituent files.

INTERFACING TO A MOLECULAR ASSEMBLER

Molecular assembler programs generate all possible constitutions compatible with the molecular formula, substructure elements, and other constraints entered by the user.[27] Their application during the structure elucidation process is extremely useful for several reasons. First, they can prove that a constitution found is unique. Second, at a given stage of the interpretation process they help to judge whether the remaining possibilities correspond to a manageable number of constitutions or not. In this second case some ordering of the constitutions generated is very helpful, especially if the list of solutions is long. Beside structure editors,[28] the ranking according to the estimated spectra is useful.[22]

C13Shift was therefore interfaced to the program ASSEMBLE.[26,28] The ^{13}C NMR spectrum of each constitution found is automatically estimated. For a comparison with the experimental spectrum, the lines are tentatively assigned such that the mean of the unsigned deviations is a minimum. The rms deviation can be used as an alternative. There is still little experience about the choice of the criterion. The constitutions are ordered according to the mean deviation per carbon atom. If the program has no parameters to estimate the chemical shift of a center, the corresponding signals are omitted.

As an illustrative example, the nine possible constitutions of the formula $C_{16}H_{34}$ which have one CH, five CH_2, and two CH_3 signals in the ^{13}C NMR spectrum were created. The results of the automatic ranking on the basis of comparisons with the experimental spectrum (39.4 d, 33.1 t, 30.1 t, 29.2 t, 26.1 t, 23.3 t, 14.2 q, and 11.0 q) are shown in Figure 5. All of the chemical shifts could be estimated by the program. The mean deviations are all small as compared with the accuracy of the model. The correct constitution is the second in the list with a mean absolute deviation of 1.12 ppm.

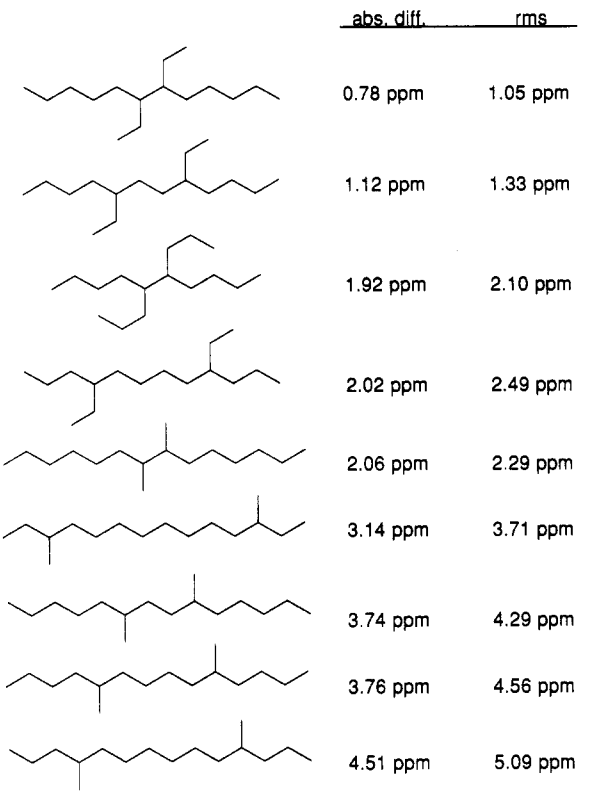

Figure 5. Ranking the structures generated by the program ASSEMBLE according to the mean absolute deviations between observed and predicted spectra. The root mean square deviations are shown for comparison.

DISCUSSION

One of the advantages of the program described in this paper is the wide range of types of carbon atoms covered by the model, i.e., its wide scope. The strategies applied to replace missing parameters by those of similar substituents or of parameters of similar substructures led to a scope of 97%, determined with a database with over 160 000 chemical shifts.[22] Without these strategies, the scope of a corresponding model[10] is estimated to about 70%.[29]

The scope of the database models essentially covers 100% of the possible query structures because of the interpolation strategy. With small databases, however, interpolation is necessary between substructures that are very dissimilar to the substructure to be estimated. Therefore, with small databases the wide scope goes along with a reduced accuracy of the prediction. The scope of the model using specific descriptors[14-18] is very much limited. It remains so, in spite of the fact that nearly 100 different models are collected into a database.[18] Tests to fit such a model to a heterogeneous set of strained alicyclic hydrocarbons showed[30] that such a widening of the scope reduces the accurracy also for this approach. Standard deviations of 3–5 ppm were obtained.[30]

Another advantage of the additivity model is the ease of extension and modification. New parameters can be added without interference with the other values. The parameters have a meaning for the chemist. In the case of the specific models[14-18] even the introduction of a single new substituent requires all model parameters to be refit. Although the extension of the database model is simple in principle (just add new data to the data file), it is not so in practice. The addition of new data and the preparation of new search files with some 100 000 compounds in the database is a time-demanding task even for mainframes.

Concerning the computational demands, the additivity model is the least demanding. However, none of the models needs highly sophisticated resources. Even the database model can be implemented on a personal computer.[31] The application of the specific model includes the relaxation of a three-dimensional structure. This might cause problems especially for noncyclic compounds.[16,17] No fully automatic structure relaxation is available that would lead to a unique final structure unbiased by the input. This is a fundamental problem for all models using three-dimensional structures for the prediction.

The main disadvantage of the additivity models is their limited accuracy as compared to the other approaches. This

seems to be an inherent problem of linear additivity models. If standard deviations better than about 5 ppm are required, other models have to be used. The models tailored specifically for selected compound classes[14-18] clearly yield the best results. When atom-centered fragments of a database are employed for the estimation,[20-24] the precision can be very good, provided appropriate references are available. However, the use of a hierarchically ordered linear list imposes some inherent limitations even if the corresponding atom-centered fragment is found for each atom of the unknown. It is not generally true that the importance of a substituent decreases with increasing distance from the center to be estimated. If, for example, the shifts of substituted benzenes have to be estimated, the model considers a NH_2- or OH-substituent in 4 position (four bonds apart) as less important than the methyl groups for a 2-isopropyl or *tert*-butyl group (three bonds apart). In reality the para OH- and NH_2-substituents contribute by −7.4 and −10.0 ppm to the chemical shift of benzene, whereas the contributions of a methyl group of a 2-isopropyl or *tert*-butyl group are only −0.6 and −1.1 ppm.[1] The model will, however, try to fit the substituents at a distance of three bonds with higher priority than those four bonds apart. If no atom-centered fragment is available, the model "interpolates" by using available entries which could be the neighboring entries in the database atom-centered fragments. This can result in very variable qualities of prediction.

Proton chemical shifts of many compound classes can be estimated with simple additivity models.[1,32] The approach used here for the estimation of carbon-13 chemical shifts can therefore be directly used for the estimation of proton chemical shifts. Concerning the accuracies, one must not forget that measurement conditions, negligible in a first approximation of ^{13}C NMR, cause variations of up to 10% of the whole chemical shift range of ^{1}H NMR spectroscopy. Estimation of coupling constants is also required if not only chemical shifts but also whole ^{1}H NMR spectra are to be predicted. Since they depend on the geometry, a structure relaxation will typically be part of a proton NMR spectrum estimation program (with all of the above-mentioned problems). Spectrum prediction furthermore implies the exact calculation of the spin system. Approximate solutions not really taking into account second-order effects[33] are of limited value only. They can lead to entirely senseless spectra which have nothing to do with reality.

ACKNOWLEDGMENT

This work was partly supported by the Swiss National Science Foundation.

REFERENCES AND NOTES

(1) Pretsch, E.; Clerc, J. T.; Seibl, J.; Simon, W. *Tables of Spectral Data for Structure Elucidation of Organic Compounds*, 2nd ed.; Springer-Verlag: Berlin, 1989.
(2) Grant, D. M.; Paul, E. G. Carbon-13 Magnetic Resonance. II. Chemical Shift Data for the Alkanes. *J. Am. Chem. Soc.* **1964**, *86*, 2984–2990.
(3) Clerc, Th.; Pretsch, E.; Sternhell, S. *^{13}C-Kernresonanzspektroskopie*; Akademische Verlagsgesellschaft: Frankfurt, 1973.
(4) Clerc, J. T.; Sommerauer, H. A Microcomputer Program Based on Additivity Rules for the Estimation of ^{13}C-NMR Chemical Shifts. *Anal. Chim. Acta* **1977**, *95*, 33–40.
(5) Cheng, H. N.; Ellingsen, S. J. Carbon-13 Nuclear Magnetic Resonance Spectral Interpretation by a Computerized Substituent Chemical Shift Method. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 197–203.
(6) Cheng, H. N.; Bennett, M. A. Trends in Shift Rules in Carbon-13 Nuclear Magnetic Resonance Spectroscopy and Computer-Aided Shift Prediction. *Anal. Chim. Acta* **1991**, *242*, 43–56.
(7) Hearmon, R. A. Microcomputer-Based ^{13}C NMR Spectral Simulation of Substituted Alkanes. *Magn. Reson. Chem.* **1986**, *24*, 995–998.
(8) Lah, L.; Tusar, M.; Zupan, J. Simulation of ^{13}C-NMR Spectra. *Tetrahedron Comput. Methodol.* **1989**, *2*, 5–15.
(9) Grasselli, M.; Olivieri, A. C. Microcomputer Program Based on the Beierbeck–Saunders Approach for the Prediction of ^{13}C-NMR Chemical Shifts in Acyclic Hydrocarbons. *Anal. Chim. Acta* **1990**, *233*, 315–319.
(10) Cf. Zupan, J.; et al. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 299–303.
(11) Fürst, A.; Pretsch, E. A Computer Program for the Prediction of ^{13}C-NMR Chemical Shifts of Organic Compounds. *Anal. Chim. Acta* **1990**, *229*, 17–25.
(12) Fürst, A.; Robien, W.; Pretsch, E. A Comprehensive Parameter Set for the Prediction of the ^{13}C-NMR Chemical Shifts of sp^3-Hybridized Carbon Atoms in Organic Compounds. *Anal. Chim. Acta* **1990**, *233*, 213–222.
(13) Pretsch, E.; Fürst, A.; Robien, W. Parameter Set for the Prediction of the ^{13}C-NMR Chemical Shifts of sp^2- and sp-Hybridized Carbon Atoms in Organic Compounds. *Anal. Chim. Acta* **1991**, *248*, 415–428.
(14) Jurs, P. C.; Egolf, D. S. Carbon-13 Nuclear Magnetic Resonance Spectrum Simulation. In *Computer-Enhanced Analytical Spectroscopy*; Meuzelaar, H. L. C., Isenhour, R. L., Eds.; Plenum Press: New York and London, 1988; pp 163–182.
(15) Small, G. W.; Jurs, P. C. Simulation of Carbon-13 Nuclear Magnetic Resonance Spectra of Cycloalkanols with Computer-Based Structural Descriptors. *Anal. Chem.* **1983**, *55*, 1128–1134.
(16) Bernassau, J. M.; Fetizon, M.; Maia, E. R. Prediction of Carbon-13 Spectra. 1. Rigid Alkanes. *J. Phys. Chem.* **1986**, *90*, 6129–6134.
(17) Bastard, J.; Bernassau, J. M.; Fetizon, M.; Maia, E. R. Computer Predictions of ^{13}C Spectra. II. Alcohols and Ketones. *Magn. Reson. Chem.* **1988**, *26*, 992–1002.
(18) Cf. Jurs, P. C.; et al. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 272–278.
(19) Ziessow, D.; Carroll, M. Referencing ^{13}C NMR Medium Shifts. *Ber. Bunsenges. Phys. Chem.* **1972**, *76*, 61–64.
(20) Bremser, W. HOSE—A Novel Substructure Code. *Anal. Chim. Acta* **1978**, *103*, 355–365.
(21) v. d. Lieth, C. W.; Seil, J.; Köhler, I.; Opferkuch, H. J. ^{13}C NMR Data Bank Techniques as Analytical Tools. *Magn. Reson. Chem.* **1985**, *23*, 1048–1055.
(22) Kalchhauser, H.; Robien, W. CSEARCH: A Computer Program for Identification of Organic Compounds and Fully Automated Assignment of Carbon-13 Nuclear Magnetic Resonance Spectra. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 103–108.
(23) Gray, N. A. B.; Nourse, J. G.; Crandell, C. W.; Smith, D. H.; Djerassi, C. Stereochemical Substructure Codes for ^{13}C Spectral Analysis. *Org. Magn. Reson.* **1981**, *15*, 375–389.
(24) Carbon-13 Nuclear Magnetic Resonance and Infrared Database, STN International, P.O. Box 2465, D-7500 Karlsruhe, Germany.
(25) Weininger, D.; Weininger, J. L. Chemical Structures and Computers. In *Comprehensive Medicinal Chemistry*; Hansch, C., Ed.; Pergamon Press: Oxford, 1990.
(26) Shelley, C. A.; Munk, M. E. CASE, a Computer Model of the Structure Elucidation Process. *Anal. Chim. Acta* **1981**, *133*, 507–516.
(27) Gray, N. A. B. *Computer-Assisted Structure Elucidation*; Wiley: New York, 1986.
(28) Shelley, C. A.; Hays, T. R.; Munk, M. E.; Roman, R. V. An Approach to Automated Partial Structure Expansion. *Anal. Chim. Acta* **1978**, *103*, 121–132.
(29) Zupan, J. Private communication.
(30) Silvestri, I. Diploma Thesis, ETH-Zurich, 1990.
(31) Gray, N. A. B. Personal communication.
(32) Cadisch, M.; Farkas, M.; Clerc, J. T.; Pretsch, E. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 286–290.
(33) Rudolf, M.; Jordis, U. Simulation of ^{1}H- and ^{13}C-NMR Spectra of Aromatic Compounds using a Personal Computer. *Chem. Intell. Lab. Sys.* **1989**, *5*, 323–327.