Journal of Computer-Aided Molecular Design, 1 (1987) 73–85
ESCOM
J-CAMD 006

WIZARD: AI in conformational analysis

Daniel P. Dolata, Andrew R. Leach and Keith Prout

Chemical Crystallography Laboratory, University of Oxford, 9 Parks Road, Oxford OX1 3PD, U.K.

Received 19 February 1987
Accepted 2 March 1987

Key words: Artificial intelligence; Conformational analysis; PROLOG

SUMMARY

A program which utilizes the techniques of Artificial Intelligence and Expert Systems to solve problems in the area of Conformational Analysis is described. The program searches conformational space in a systematic fashion, based on the technique known as heuristic state-space search. The program proceeds by recognizing *conformational units*, assigning one or more *conformational templates* to each unit, and *joining* them to form conformational suggestions. These suggestions are *criticized* to discover logical inconsistencies, and any resulting stresses are *resolved*. The resulting conformational suggestions are sometimes accurate enough for immediate use, or may be further refined by a numerical program. The latter combination is shown to be quite efficient compared to purely numerical conformational search techniques.

INTRODUCTION

The physical properties, biological activities and reactions of organic molecules are intimately related to their conformation(s). Without a good understanding of these conformation(s), we cannot hope to fully predict such properties. To better appreciate the behavior of a flexible molecule, it is of obvious importance to be able to examine *all* of its low energy conformations. To this end, many programs have been written in the last few years. The majority of these are numerically based, utilizing techniques such as quantum mechanics [1], molecular mechanics [2], and/or distance geometry [3]. Quantum mechanical programs can be highly accurate, but require enormous amounts of CPU time for large molecules. Molecular mechanics programs can give excellent results, but require careful parameterization. They may also require large amounts of CPU time when the conformational space of a flexible molecule is studied by torsional angle driving, Monte Carlo search techniques, or molecular dynamics. Distance geometry requires both significant CPU power and a good set of distance constraints, generally based on experimental data. Recently a few programs have been written which utilize models and/or symbolic knowledge [4]. Their purpose is to provide acceptable answers in substantially shorter times than the numerical programs. However, none of them completely fulfil the chemist's requirements. Some of them have a limited domain, others use over-simplified knowledge, or use data which may be subtly prejudiced to predict just a single conformation. None of these programs try to account for the intra- and

0920-654X/87/$ 3.50

inter-molecular forces that may affect a molecule's conformation(s), without recourse to strictly numerical methods.

Our research examines the utilization of applied Artificial Intelligence techniques in examining conformational space. Previously, we have described our pilot project, WIZARD-I, which could predict the conformational spaces available for acyclic molecules [5]. The success of that small-scale study gave us confidence to embark on a project of much broader scope. We wish to introduce the result of this project, a program named WIZARD-II, which applies these techniques to a much wider range of compounds, including (to date) acyclic and cyclic structures, unsaturation and heteroatoms, and lays the groundwork for oligosaccharides, peptides, organo-metallics and inorganic structures. In this paper we will give an overview of how the program works, and introduce the fundamental concepts and algorithms of this approach. Later papers will expand upon the details of these algorithms.

DISCUSSION

The applied Artifical Intelligence (AI) techniques which we explored in WIZARD are called Expert System (ES) techniques, because they attempt to reproduce in some fashion the behavior of an expert in a given problem domain [6]. We chose to investigate this approach because we observed that an expert in conformational analysis could significantly help the process of exploring a molecule's conformational space. We noted that in the case of simple molecules (e.g. acyclic drugs and insect pheromones) an expert could make good enough 'guesses' of the geometries of local minima, such that after minimization by a molecular or quantum mechanics program the initial and minimized conformations would be recognizably the same (our 'sameness' criterion is less than 0.25 Å RMS difference). In many cases, the process of guessing and minimization proved to be much more efficient than common numerical methods. The advantages of this approach compared to straight torsion angle driving lie in the fact that only guesses which an *expert* has approved are minimized. In addition, the expert's starting geometries lie close to the actual minima, so that the minimizations converge rapidly. The obvious pitfalls are: some minima might be overlooked, the generation of a large number of guesses might be too time consuming, and that conformational analysis experts are not commonly available. Many of these problems could be obviated by using an 'artificial expert' (*i.e.* an ES). An ES could search conformational space rapidly and systematically, thus reducing the chances that local minima would be missed. These guesses would then be minimized by a molecular or quantum mechanics program, and the bad ones thrown out. This ES would be available as a stand-alone computer program, thus providing chemists with an 'in-house' conformational analyst. This ES will be successful if the combined computer time for this process of guessing and minimizing is less than the time required to obtain the same results by utilizing the minimization program with a numerically-based conformational search (torsion angle driving, Monte Carlo, *etc.*), or if it returns information that is not available with current techniques (such as flexibility, solvent dependence, or conformations for non-parameterized systems).

WIZARD works as an intermediate between a chemist's input and a conformational analysis program, such as MM2. WIZARD is not a molecular graphics program *per se*, but it can utilize input from a number of different graphics programs such as MIMIC [7], COSMIC [8], etc., and can call on these programs directly to display results. In addition, WIZARD does not actually

contain a molecular mechanics (MM) program, but it can call a number of MM programs directly for various purposes (refinement of results or intermediate guesses, template building, etc.). Once the molecule is input by one of these other stand-alone programs, the user then runs WIZARD and tells it the name of the molecule file. WIZARD then proceeds to perform a conformational analysis of the molecule. The approach adopted in WIZARD has the following structure:

The molecule, defined by atom type and connectivity, is input
The molecule is analyzed, and *conformational units* are recognized
An *abstract representation* of the molecule is generated
Conformational Space is heuristically searched
    The next conformation is *suggested* for the abstract representation
        Each unit is assigned the next *subconformation*
    This suggestion is *criticized* to see if problems exist
    If no problems exist, the suggestion is *built* from coordinate *templates*
        The templates for each subconformation are *joined*
        Each assembly step is *criticized*
    The entire conformation is *criticized*
    Critical problems are *resolved*
    A three-dimensional model of each suggestion is output
Until the heuristic search is complete

The first step is to recognize various features of the molecule. WIZARD accepts molecules in a number of different file formats. These are all converted into a single internal format, similar to that used in the MIMIC programs [7]. Since WIZARD is capable of using crude two-dimensional sketches of the molecule, the stereochemistry of each chiral atom is checked, and if it has not been specified the user is asked to define it. Next, the topological symmetry of the molecule is discovered using a variation of the Gasteiger method [9]. This method suggested itself very strongly to us since it is based on a clearly defined set of recursive rules, which were easy to translate into PROLOG code. These rules have since been translated into FORTRAN for the sake of efficiency. Following this, WIZARD recognizes the *conformational units* comprising the molecule, as shown in Fig. 1. A conformational unit is defined as a connected set of atoms and bonds about whose conformational behavior the program has knowledge. For example, there is a substantial body of knowledge about the conformational behavior of cyclohexane rings. Some units are contained within other units, e.g. cyclohexane is contained within decalin. The units are arranged in a hierarchical fashion so that the more complex units are utilized in lieu of simpler units, unless later analysis shows that no reasonable conformations can be constructed utilizing these complex units. In that case, the information contained in the simpler units will be utilized to try to find suitable conformations.

Some of the information contained in the unit knowledge base is shown in Fig. 2. Cyclic units consist of one or more rings, which may be fused in any fashion. Acyclic units generally consist of from one to three bonds, and optionally, an environment consisting of the neighboring shells of atoms. Atom and bond types may be specified exactly, or as one of a number of increasingly more general classes. Stereochemistry may also be specified if desired. In addition to the pattern of atoms and bonds which identify the unit, the frame contains a series of *templates (vida infra)* which contain the knowledge about specific conformations and their energies, coordinates, flexibilities,

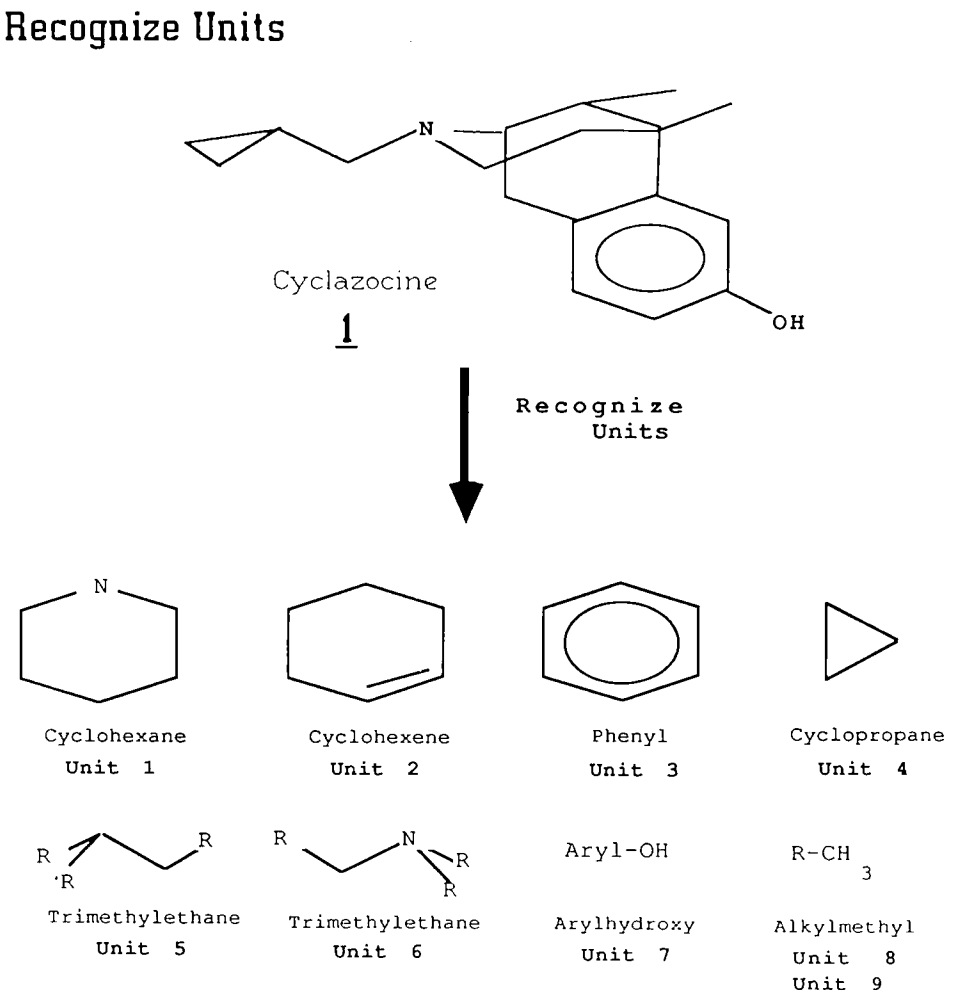


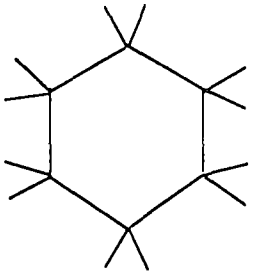
Fig. 1. Recognition of conformational units in cyclazocine.

etc. Once these units have been recognized, an abstract model of the molecule is built up as shown in Fig. 3. In the case of cyclazocine **1**, WIZARD finds nine units. The connections of these nine units are used to derive a unit graph of the molecule, as shown in the figure. By recognizing certain features in this abstract graph, the most efficient search can be derived. WIZARD currently uses a modification of the A* algorithm [20], where the heuristic values are based on estimated strain and criticism derived from past experience. The search tree mirrors the structure of the abstract graph, and the starting node is chosen to be that unit which is most highly constrained (which we call the *crucial unit*). The crucial unit is defined as the most highly connected unit. In the case when two units have equal number of connections, the one with the most complex set of connections is more crucial. Ring bridging is considered to be more complex than ring fusion, which is more complex than acyclic connections.

Unit Knowledge

Pattern:

- Atom Types
- Bond Types
- Stereochemistry



Template List:

- Base Strain Energy
- X, Y, Z
- Abstract Bond Angles (s, ns, ne, e)
- Ring Shapes (planar, U, S, etc)
- Symmetry Information
- Flexibility

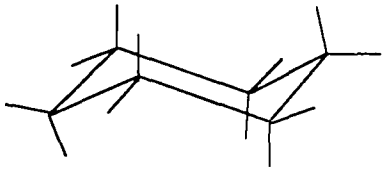
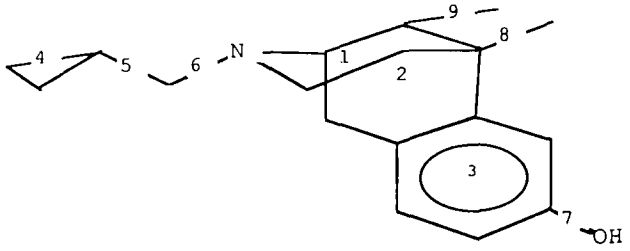


Fig. 2. Types of information stored in a unit frame.

Hierarchical Description



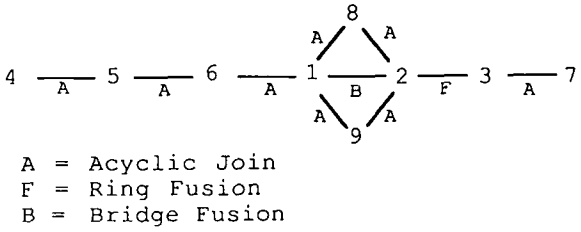


Fig. 3. Abstract description of cyclazocine.

Chemist Supplied Critics (Long Term Memory)

Energetics	Are the template energies too high?
Join	Are we trying to join incompatible bonds? e.g. fusing an eclipsed and a staggered bond
VDW	Is this a known problem case? e.g. butane g^+ and butane g^- (pentane violation)

Self Learned Critics (Short Term Memory, from previous tries during this analysis)

bad_join	[unit1, chair, b123][unit2, halfchair, b231]
bad_vdw	[unit1, chair, b234][unit6, sp3sp3_3_gp, b1][unit5, sp3sp3_3_gm,b1][unit4, c3planer,b1]

Fig. 4. Types of abstract critics utilized.

This abstract graph is then examined to find higher level entities, such as *extended ring* systems, and *super rings*. An extended ring system is defined as a linearly fused or bridged path of rings. A super ring is defined as three or more rings which are all commonly joined. The recognition of these higher level entities is important for more advanced reasoning techniques that adjust the molecule to account for intra and inter molecular effects (for example, to relieve strain, fit desired geometries, etc.). Changing a ring in an extended system may require the adjustment of adjoining rings (and so on in a recursive fashion), while changing a ring in a super ring may cycle back and require further adjustment of that first ring, or prove impossible (consider trying to deform a chair cyclohexane in adamantane (super ring) as opposed to cholestane (extended ring)). Note that in cyclazocine, the cycles 1-2-8 and 1-2-9 are not considered super rings because units 8 and 9 are not rings.

After the recognition step, WIZARD then attempts to find viable conformations by performing a systematic search in conformational space. WIZARD performs this search in a hierarchical fashion. A symbolic (as opposed to numeric) suggestion is built at the abstract level, which is then criticized for problems. The types of criticism which can be applied at the abstract stage are shown in Fig. 4. The program looks for connections of units which are historically known to be bad for various reasons — for example, a C_5 acyclic chain containing succesive gauche$^+$ gauche$^-$ C_4 units is known to be bad for van der Waals reasons (the so-called *pentane rule*). We can also examine the suggestion to see if something similar has already been tried and rejected. In complex molecules these heuristics can reduce the CPU time needed to perform an exhaustive search by as much as an order of magnitude. Other abstract critics based on abstract geometry and grammar theory are under investigation.

If the symbolic suggestion passes these abstract critics, it is then built up in a stepwise fashion from coordinate templates, with more criticism at each step. These templates can be obtained in several different fashions. The two we currently use are generation by molecular mechanics and extraction from the Cambridge Structural Database (CSD). These two methods each have their own advantages and drawbacks.

Searching the CSD can be rapid, given the right algorithms [10], but even if the entire database is used, it is possible that certain classes of compounds might not be present. And if they are, it is also possible that only a small set of the allowable low energy conformations might be represented. In addition, without the ability to reason about strain inducing factors, such as van der Waals repulsions, electrostatic forces, *etc.*, it is difficult to decide which conformations correspond to appropriate minima. This problem has been addressed by statistical means [11] but this approach does not tell us which minima might be appropriate for the types of strains present in a given molecule. In addition, there are subtle biases in any crystal database due to entropic effects of crystallization. Compounds with multiple low energy conformations seem to be more difficult to crystallize than compounds with a single well-defined low energy form, probably due to the higher entropic penalty. Thus, this class of compounds will be under-represented in a crystal database, both as per total number of crystals prepared, and as regards sampling the total number of conformations available to each compound.

Utilizing systematic search techniques in molecular mechanics to find conformational minima can be much more time consuming, especially if a complete search of the conformational space is desired, but this can be offset by the fact that we only need perform such an analysis once for each unit pattern. It is also possible to have WIZARD generate templates by creating crude suggestions which are then refined using MM2 to create the corresponding energy minimized geometry. WIZARD then processes each suggestion to eliminate symmetrically equivalent suggestions, and to create the template files. This was done with several steroid skeletons, for example. This process forms part of the foundation of a learning machine, which is currently under investigation.

The first assemblage is built by taking the appropriate template for the crucial unit, and *joining* the template for the next most crucial unit. The exact details of the joining algorithm will be presented in a future paper. The join is criticized depending on the quality of fit between common atoms and bonds (several weighting schemes may be utilized). The joined assemblage is then criticized on the basis of van der Waals forces, and optionally, on other intramolecular forces. The assemblage can be minimized before proceeding, but this is currently not done so as to test more rigorously the limitations of the logical method without numerical help. If the assemblage is criticized, the minimal path necessary for failure is deduced based on this and previous failures, and saved to provide abstract critics for use in short-term memory (lasting for only this run) and eventually for the learning machine which will provide long-term critics for all future runs. If the assemblage passes this barrage of criticism, the next unit template is joined and the criticism is repeated, until the entire molecule is constructed. It is then announced to the user, optionally displayed, and saved for later usage or minimization. WIZARD then proceeds with the systematic search of conformational space until all possible suggestions have been considered. The utility of this approach can be seen by examining Fig. 5, which shows one of WIZARD's suggestions for cyclazocine overlaid upon the X-ray crystal structure [12], Fig. 6, which shows one of WIZARD's suggestions for cholestane overlaid upon the crystal structure for dichloro cholestane [13] and Fig. 7 which compares one of WIZARD's suggestions with the crystal data for dextromethorphan [14]. We will discuss the CPU time required for these analyses later in this paper.

However, not all molecules have low energy conformations which can be built by such an approach. A simple example is flufenamic acid (**2**), which is a strained molecule. If one tries to build any conformation of **2** from standard anilinyl templates which require co-planarity of the phenyl ring and the NR_2 substituents, a pair of hydrogens will be compelled to lie in the same volume of

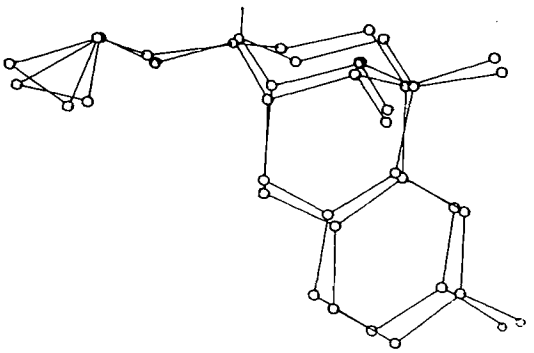


Fig. 5. Cyclazocine: Overlay of WIZARD's suggestion with the X-ray crystal structure.

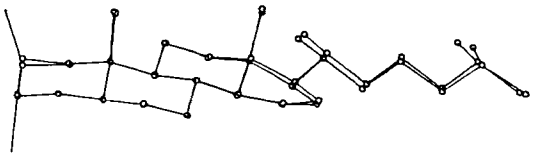


Fig. 6. Cholestane: Overlay of WIZARD's suggestion with the X-ray crystal structure.

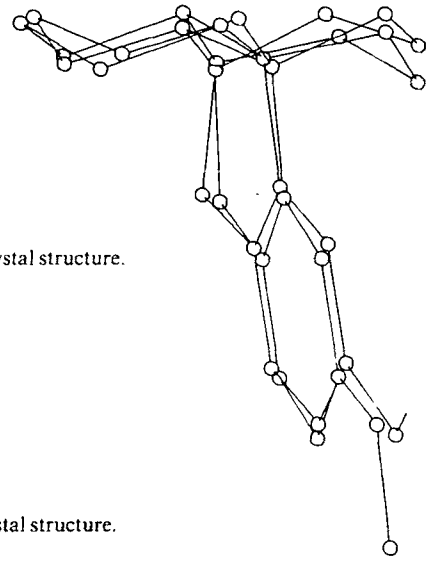
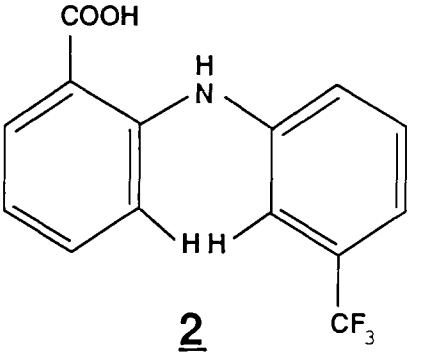


Fig. 7. Dextromethorphan: Overlay of WIZARD's suggestion with the X-ray crystal structure.

space (Fig. 8). Thus, all simple conformational suggestions for **2** are criticized. The resolution of such criticism is an exciting problem. This resolution is performed in an iterative fashion. First, a naive suggestion for strain relief is proposed for each criticized pathway. The possible suggestions are ranked in order of energy, with torsion angle adjustment being preferred before bond angle changes, which are in turn more favored than bond length changes. This combination of all suggestions is then criticized, to see if there are any logical, geometrical or chemical inconsistancies in them. For example, it may arise that two separate 'violations' suggest twisting the same bond in a positive and negative direction simultaneously. In such cases, WIZARD would eliminate this



2

Fig. 8. One possible planar structure for flufenamic acid.

suggestion at the abstract stage. Those suggestions which are internally consistant are then created, either by manipulating the criticized structure or re-creating it from previously deformed templates. These deformed conformations are then recriticized to see if any problems due to intramolecular forces remain or have arisen. If the deformed conformation is not criticized, WIZARD then outputs this for the user to examine. If all of the resolution suggestions are eliminated at the abstract stage, or the deformed conformations all exhibit problems, WIZARD then decides that low energy paths are probably incapable of relieving all strains in the molecule, and so considers higher energy processes as well.

However, in flufenamic acid, the strain can be resolved by simple torsion angle changes. WIZARD suggests four resolved suggestions. There are two known crystal morphs [15] for flufenamic acid, and these are shown overlaid on WIZARD's corresponding suggestions in Fig. 9 and 10. In addition to resolving strain brought about by van der Waals repulsion, WIZARD can also resolve strain due to electrostatic and hydrogen bonding forces. For example, the torsion angle of the methoxy carbonyl sidechain of cocaine is strongly dependent on the degree of polarization of the quaternary nitrogen and the resulting attraction of the oxygen lone pair electrons [16]. A comparison between WIZARD's initial suggestion which disregarded such effects and the crystal structure of cocaine hydrochloride [17] is shown in Fig. 11. The resulting suggestion given by WIZARD after the strain arising from such forces is resolved is shown compared to the crystal data in Fig. 12.

In di–*t*–butylmethane, no combination of one or more lower energy torsion angle twisting processes is capable of resolving all of the van der Waals repulsive strain. In this case, WIZARD suggests that the higher energy process of opening the central angle about 15° must also be used, and this is in agreement with electron diffraction data reported in the literature [18]. This process has been extended to the selection and manipulation of ring templates as well. The [3.3.1] bicyclononane skeleton shows substantial repulsion between hydrogens in the 3- and 4-positions, requiring some strain relief. A comparison between WIZARD's resolved suggestion for 9-amino

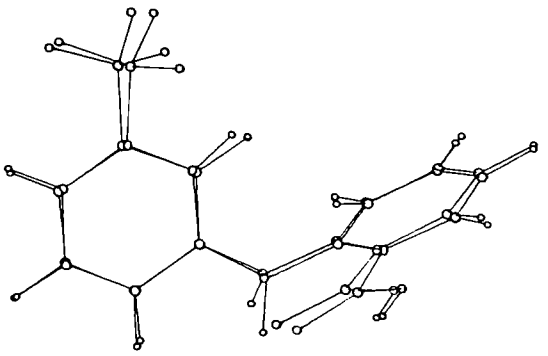


Fig. 9. Flufenamic acid: Overlay of one of WIZARD's suggestions with the X-ray crystal structure of one reported crystal morph.

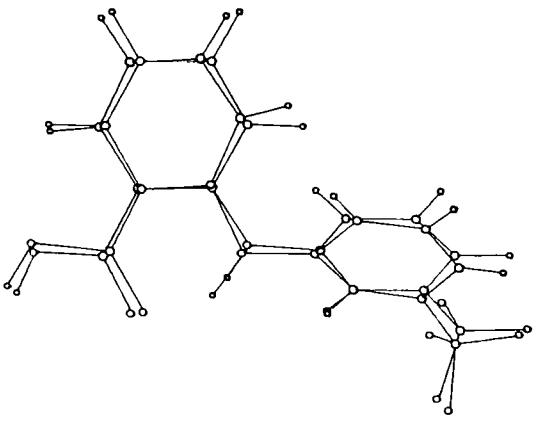


Fig. 10. Flufenamic acid: Overlay of another one of WIZARD's suggestions with the X-ray crystal structure of a different crystal morph.

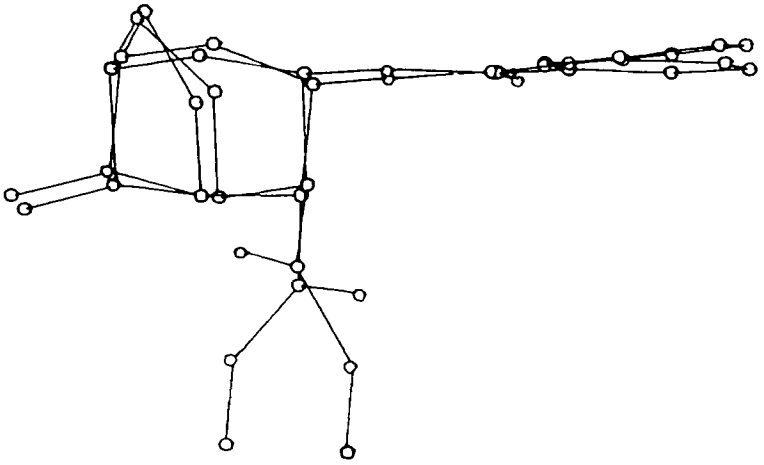


Fig. 11. Cocaine hydrochloride: Overlay of the X-ray crystal data with WIZARD's suggestion prior to strain resolution.

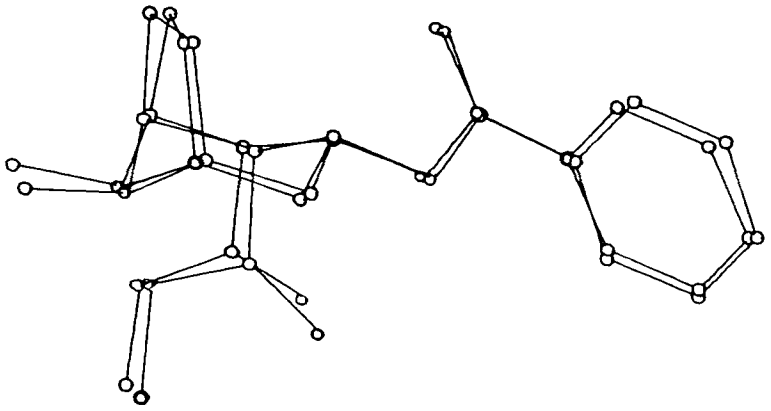


Fig. 12. Cocaine hydrochloride: Overlay of the X-ray crystal data with WIZARD's suggestion after strain resolution.

[3.3.1] bicyclononane 9-carboxylic acid and the crystal structure [19] is shown in Fig. 13. In this case, a combination of ring twisting and flattening was required to effect the necessary separation between the offending '*trans*-annular' hydrogens.

CONCLUSIONS

Some of the criteria for sucess that we had laid down for the performance of a conformational analysis ES were that the results must be accurate, and that the CPU time required to return these results must be less than that needed to obtain equivalent results utilizing availiable numerical techniques. It can be seen by examining the figures that WIZARD's results are quite close to the crystal data. For many purposes these results would suffice without further refinement by numeri-

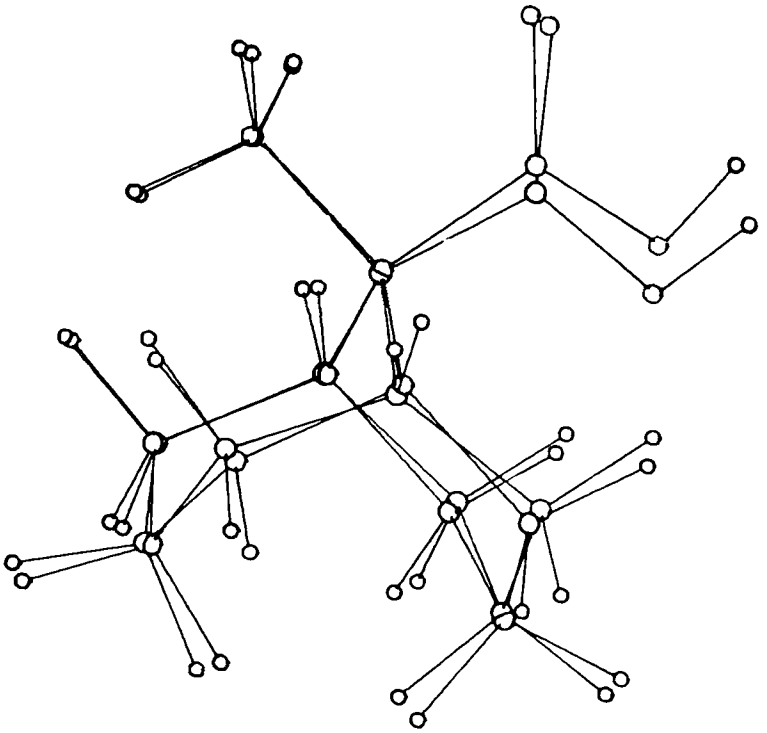


Fig. 13. 9-Amino [3.3.1] bicyclononane 9-carboxylic acid: Overlay of the X-ray crystal data with WIZARD's suggestion after strain resolution.

cal programs. However, WIZARD currently cannot estimate relative energies better than $\pm 25\%$, so that investigations requiring good strain energy values would still require such refinement. It is important to assure ourselves that utilizing WIZARD as a pre-screen is not simply a waste of computer resources when we have to utilize these numerical programs anyway.

The question of efficiency can be answered by looking at the relative CPU times required by WIZARD and by MM2 for various molecules. While running on a VAX 11/750 (see details below), WIZARD required 681 CPU seconds to generate 15 conformations for cyclazocine (about 45 CPU seconds each), while MM2 required an average of 401 CPU seconds to minimize each one. The time required for WIZARD to generate 16 suggested local minima for cocaine was 1167 CPU seconds (about 73 CPU seconds each), compared with MM2's average requirement of 351 CPU seconds to minimize each of these suggestions. It is fair to say that using WIZARD as a pre-screen costs less CPU time than required to minimize two or three bad guesses from another conformation generator (for example, either Monte Carlo methods or the torsion angle driver). To obtain the same results with a pure torsion angle driver, one would need to drive two bonds on the side chain by three (or more) steps each, and also to drive about several bonds in the non-aromatic rings. It is hard to estimate exactly how many steps would be required to search the conformational space available to the rings by torsion angle driving, but it is almost certainly an underestimate to say that driving two bonds by at least three steps would be required. It is well

known that there are pitfalls in the use of the torsion angle driving method for the calculation of conformational interconversions in rings [21], so that it might be difficult to obtain all ring conformations without utilizing even more steps. Thus, it is likely that utilizing the torsion angle driving method would require more than 3^4 steps, each requiring roughly 400 CPU seconds to perform the associated minimization. It can be seen that is certainly more efficient to utilize WIZARD as a pre-screen, and then to utilize MM2 to refine the suggestions of this Expert System.

It is also important to consider the scope and limitations of the process of trying to build models by joining templates, since many programs are now being written which behave in this fashion. We have shown that this technique works well in certain cases, but fails in others. Molecules which do not have intramolcular forces deforming the templates can be built quite well without any further analysis (e.g. cyclazocine, dextromethorphan and cholestane). But if there are intra- or intermolecular forces which act across template boundaries, then that technique provides naive answers at best, and quite erroneous answers at worst. Relying upon a molecular mechanics algorithm to resolve all of the problems will probably provide some conformational minima, although much of the conformational space may go undiscovered. For example, if a program were to build an all-planer model of flufenamic acid from anilinyl templates, and then minimize this suggestion, the user would obtain a single result from each starting conformation. However, as we have shown in the WIZARD-I paper, strain relief often leads to a process we call 'conformational splitting', leading to two or more local minima for each critical strain path. In flufenamic acid the bond which is twisted can do so in either a positive or negative fashion, leading to two conformers for each planar starting conformation. Thus is can be seen that this sort of effect must be accounted for in any process that attempts to search conformational space. This is done during Monte Carlo or torsion angle driving techniques by choosing a large number of randomizations or sufficiently small step increments. We have shown that this analysis can also be performed by a technique based in symbolic logic, with a high degree of efficiency.

IMPLEMENTATION

WIZARD is written using a combination of PROLOG and FORTRAN-77, and runs on a VAX 11/750 under VMS 4.3. The Expert System, comprising the naive suggestors, critics, arithmetic and geometric reasoning, and learning portions of the program are written in PROLOG, while the molecular input and output, recognition, joining, and force determination portions are written in FORTRAN. The PROLOG currently used is part of the University of Sussex POPLOG package, but WIZARD maintains Clocksin and Mellish [20] compatibility to facilitate ease of conversion to other PROLOG systems.

ACKNOWLEDGEMENTS

The authors thank the Science and Engineering Research Council (U.K.) and Smith Kline and French Research Ltd. for financial support. DPD thanks Docent Robert E. Carter and Prof. Börje Wickberg of the University of Lund and the Swedish Board for Technical Development (STU) for financial support and encouragement for the WIZARD-I pilot project.

REFERENCES

1 *a*. Singh U.C. and Kollman, P., QCPE Bulletin, 2 (1982) 117–145. (Gaussian 80)
b. Dewar, M.J.S. and Thiel, W., J. Am. Chem. Soc., 99 (1977) 4899–4917 (MNDO)
c. Pople, J.A. and Beveridge, D.L., J. Chem. Phys., 47 (1967) 2026–2033. (CNDO)

2 *a*. Wiberg, K.B. and Boyd, R.H., J. Am. Chem. Soc., 94 (1972) 8426–8430.
b. Burkert, U. and Allinger, N.L., A.C.S Monograph, 177 (1982) 72–114.

3 Crippen, G.M., Distance Geometry and Conformational Calculations, Chemometrics Research Studies Series, 1, Research Studies Press, Wiley, 1983.

4 *a*. Corey, E.J. and Feiner, N.F., J. Org. Chem., 45 (1980) 757–764. (A part of LHASA limited to cyclohexane rings only.)
b. Cohen, N., Colin, P. and Lemoine, G., Tetrahedron 37 (1981) 1711–1721. (An excellent program, which joins simple templates with some abstraction, but does not examine intramolecular forces and strain relief directly.)
c. De Clercq, P.J., Tetrahedron, 40 (1984) 3717–3738. (A BASIC program limited to joining ring templates.)
d. Wipke, W.T. and Hahn, M., ACS Symposium Series, 306 (1986) 136–146. (A template assembly program which obtains its templates from an abbreviated X-ray data base.)

5 Dolata, D.P. and Carter, R.E., J. Chem. Inf. Comput. Sci. (in press).

6 Barr, A. and Feigenbaum, E. (Eds) The Handbook of Artificial Intelligence, Vol 2., William Kaufmann, Inc., Los Altos, CA, 1982.

7 Liljefors, T., J. Mol. Graph., 1 (1983) 111–117.

8 Vinter, J.G., Davis, A. and Saunders, M., J. Comput.-Aided Mol. Design, 1 (1987) 31–51.

9 Jochum, C. and Gasteiger, J., J. Chem. Inf. Comput. Sci., 17 (1977) 113–117.

10 Allen, F.H., Bellard, S., Brice, M.D., Cartwright, B.A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B.G., Kennard, O., Motherwell, W.D.S., Rodgers, J.R. and Watson, D.G., Acta Cryst., B35 (1979) 2331–2339.

11 *a*. Murray-Rust, P. and Motherwell, S., Acta Cryst., B34 (1978) 2534–2546.
b. Murray-Rust, P. and Raftery, J., J. Mol. Graph, 3 (1985) 50–56.
c. Taylor, R., J. Mol. Graph., 4 (1986) 123–131.

12 Karle, I.L., Gilardi, R.D., Fratini, A.V. and Karle, J., Acta Cryst., B25 (1969) 1469–1479.

13 Geise, H.J., Romers, L. and Rutten, E.W.M., Acta Cryst., B20 (1966) 249–257.

14 Gylbert, L. and Carlström, D., Acta Cryst., B33 (1977)2833–2837.

15 *a*. McConnell, J.F., Cryst. Struct. Comm., 2 (1973) 459–461.
b. KrishnaMurthy, H.M., Bhat, T.N. and Vijayan, M., Acta Cryst., B38 (1982) 315–317.

16 Shen, M., Ruble, J.R. and Hite, G., Acta Cryst., B31 (1975) 2706–2709.

17 Gabe, E.J. and Barnes, W.H., Acta Cryst., 16 (1963) 796–801.

18 Bartell, L.S. and Bradford, W.F., J. Mol. Struct., 37 (1977) 113–126.

19 Bhattacharjee, S.K., Chacko, K.K. and Zand, R., Acta Cryst., B35 (1979) 399–401.

20 Clocksin, W.F. and Mellish, C.S., Programming in PROLOG, Springer-Verlag, Berlin, 1981.

21 Burkert, U. and Allinger, N.L., J. Comp. Chem., 3 (1982) 40–46.