Journal of Computer-Aided Molecular Design, 6 (1992) 385–396
ESCOM

J-CAMD 165

Automated site-directed drug design: The generation of a basic set of fragments to be used for automated structure assembly

P.-L. Chau and P.M. Dean*

Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QJ, U.K.

Received 10 March 1992
Accepted 31 May 1992

Key words. Drug design; Structure assembly; Molecular fragments

SUMMARY

If a method is to be developed to assemble putative ligand structures in site-directed drug design, from molecular graphs generated in the site, then basic building blocks are needed. Structure assembly is a combinatoric process that needs to be optimised if it is to be tractable. What has to be determined is whether small molecular fragments can have transferable properties from one molecule to another. In this paper we determine all possible combinations of 3-, 4- and 5-atom aliphatic fragments from a small set of atoms H, C, N, O, F or Cl. The frequency of occurrence of these candidate fragments is searched for in the Cambridge Structural Database. A similar analysis is performed on charged fragments. A more restricted search is carried out for P and S and aromatic structures. A basic set of fragments can be derived that have a significant frequency in known crystal structures. The transferability of fragment properties is discussed in subsequent papers.

INTRODUCTION

The work described in this and subsequent papers [1,2] tries to address the problem of whether small molecular fragments, with transferable properties, can be generated for further use in automated site-directed drug design. These three papers form a coherent unit and readers may like to consider them together rather than in isolation. If the crystal structure of the site is known, and the position of the site atoms can be defined to an adequate resolution, then the properties associated with the atoms in the site should provide all the constraints for building novel molecules. Can small molecular fragments, rather than a sequence of single atoms, be used to build novel ligands within the constraints imposed by the site?

In computer-aided drug design, the operator uses structural knowledge, chemical information

* To whom correspondence should be addressed.

0920-654X/$ 5.00

and experience to create a structure with the appropriate properties to fit the site. The question we seek to resolve is: how much of this work can be automated? There are three conceptual problems that have to be explored. Firstly, the site has to be characterised; it needs to be surveyed to establish the spatial pattern of molecular forces which would act as constraints in the generation of a putative ligand structure. Secondly, a molecular graph has to be developed to fit, or span the site. Thirdly, atoms have to be placed on the molecular graph subject to the constraints of valency rules and the molecular forces generated by the site. These, of course, are only conceptual steps; it may be possible to generate a ligand in a site by analytical methods without needing to divide the process into these separate parts. However, for the moment, a stepwise approach has the advantage that each step can be explored and checked for unforeseen factors which might otherwise frustrate the search for a better method.

Sites can be readily surveyed for electrostatic potential and hydrogen-bonding regions [3,4]. The GRID program [5,6] can be used to investigate the interaction of small molecular probes with the site to locate where the probe interaction would be favourable. Generation of a molecular graph can be achieved using libraries of spacer skeletons [7,8] or through the docking program using structures from a shape database [9,10]; chain structures can be grown over macromolecular surfaces to link molecular subgraphs within a site [11]. The ALADDIN program [12] uses a molecular database and tries to extract substructural features which may then be used to piece together a new putative ligand using parts from a chemical database. No attempt is made in ALADDIN to confront the combinatorial complexities of structure generation and placement.

In this series of papers we explore whether small molecular fragments may be used as building blocks for eventual use in an automated molecular structure generator. Suppose that we have a molecular graph with N vertices and that we can choose atoms from j chemical elements, then there are j^N possible atom placements on the graph. Many possibilities would be forbidden by the valence rules for adjacent atom types. The difficulty with attempting to place the atoms on the graph, one atom at a time, is that the atom properties may not be directly transferable. This obstacle would necessitate the calculation of the molecular atomic properties, by some molecular orbital procedure, each time a putative ligand is proposed. For a combinatorial process, proportional to j^N, even excluding forbidden structures, continual recalculation would rapidly become impracticable. Furthermore, some atomic arrangements, although possible from valence rules, may be highly improbable. Thus it may be better to consider the atoms as belonging to groups of 3-, 4- or 5-atom fragments provided that it can be shown that the fragments exist and their properties are transferable.

The fragments can be generated combinatorially for all atoms, from a specified set, and their frequency of occurrence in the Cambridge Structural Database (CSD) can be found. Geometric, electronic, hydrogen-bonding and hydrophobic properties, where appropriate, can be assigned to the fragments. The transferability of these properties can be tested by comparing those calculated from the fragment with the properties found in a statistically large set taken from the CSD. Dissection of molecules into fragments is not a new concept; molecular hydrophobicity has been calculated in this way by Rekker [13] and by Hansch and Leo [14]. Clementi [15] also proposed that molecular interactions could be studied similarly.

Complexity in site-specific structure generation

Site-specific structure generation is a very complex problem. For the simplest model, consider

a receptor site with a number of site points. These site points should interact with ligand points on the generated structure. Typically, the site points could be hydrogen-bonding groups, atoms with a high residual electronic charge or strong hydrophobic groups. The geometry and chemical properties of the other atoms in the site could also vary by a great amount. Moreover, the interaction between atomic groups at the site points and those at the ligand points is not necessarily a one-to-one relationship. An atomic group at a ligand point could interact with two atomic groups belonging to different site points of the receptor and vice versa. In more complex systems, co-ligands such as co-enzymes and water molecules could modulate the ligand–receptor interaction.

In this work, fragments, as opposed to atoms, are tested for use as the basic building blocks for drug design. In many cases the atomic residual charges of the atoms in the fragment have been found to be transferable. For example, the residual electronic charge of an oxygen atom can vary over a large range, from about -0.5 to 0.0 units of an electron charge, but the charge shows narrower statistical limits for certain atomic arrangements. In the design of ligands, if we know that a certain ligand point requires an atom with a residual electronic charge of, say -0.2, we might want to assign a particular class of oxygen atom to that ligand point. However, unless we know the neighbouring atoms and their bonding patterns around that oxygen atom, it is impossible to predict the charge on the oxygen. Fragments, however, as will become apparent from these papers, generally have more predictable properties. For example, if we know that the ethanolic oxygen has a residual electronic charge of about -0.2, then other aliphatic hydroxyl oxygen atoms probably have similar charges. We could assign a hydroxyl oxygen to the ligand point as one possible atom placement. In this way one could gradually build up an optimum ligand.

COMPUTING METHODS AND RESULTS

Fragment generation

This section describes how chemical fragments, composed of 3, 4 or 5 atoms from a specified set of atoms, are generated exhaustively and which fragments are commonly occurring. The method ensures that all geometrically possibly chemical fragments are generated, and that no human bias is introduced into the choice of fragments. These fragments are later tested for their suitability as basic building blocks for drug design.

We require a structure generator that takes a specified set of atoms, goes through all possible combinations of bond types with all possible atom types, and outputs all possible fragments containing 3, 4 or 5 of these atoms. This structure generator should exhaustively generate all fragments. No duplicate structures should be output either. Thus, we can be certain that all possible chemical structures are considered.

Generation of aliphatic fragments with STRUCGEN

The structure generator in this work, STRUCGEN, generates all chemical fragments containing 3–5 atoms. The atoms used are from the set (carbon, nitrogen, oxygen, hydrogen, and X, where X is either fluorine or chlorine). All halogen compounds contain either fluorine *or* chlorine, and none of them contains fluorine *and* chlorine. This is because it is generally more difficult to synthesize mixed halogen compounds, so they are not considered. All fragments contain a central atom, which has its valencies filled, and two to four peripheral atoms, which may or may not have their valencies filled. Delocalised, aromatic or cyclic structures are not generated by this programme.

Some of these structures are handled separately at a later stage.

STRUCGEN takes no input. It outputs the atomic number of the atoms of the fragment in a fixed order, starting from the central atom, and also the adjacency matrix of the equivalent chemical graph of the fragment. The formal bond order (single, double or triple) between adjacent atoms is specified by the entry in the adjacency matrix, **A**. An entry of zero at *ij* means that the *i*-th atom is not bonded to the *j*-th atom. Note that this matrix is symmetric. All entries in the diagonal ($i = j$) are used for storing information about the type of element at the *i*-th position.

In its internal memory, STRUCGEN has the information regarding bond types allowed for each atom (triple bonds for carbon and nitrogen, double bonds for carbon, nitrogen and oxygen, and single bonds for carbon, nitrogen, oxygen, hydrogen, fluorine and chlorine), and the valencies for each atom. The programme takes an atom and adds different atoms to it until the valencies of the central one are fully occupied. Diatomic entities are rejected because they are too small to be of use as fragments. All molecules produced by STRUCGEN are also rejected because they do not contain empty valencies for attachment to other chemical moieties. If this attachment is not possible, these molecules cannot be used as parts of larger molecules.

The fragments are specified by the adjacency matrices, whose diagonal entries dictate the elements at different positions, and whose off-diagonal elements denote the connectivity of the equivalent chemical graph. However, fragments that are represented by different adjacency matrices may be identical because of symmetry properties of the equivalent chemical graphs. Consider the following two fragments that can be denoted in two equivalent ways:

$$A_1 = \begin{pmatrix} 6 & 1 & 1 & 2 \\ 1 & 6 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 7 \end{pmatrix} \quad A_2 = \begin{pmatrix} 6 & 2 & 1 & 1 \\ 2 & 7 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 6 \end{pmatrix}$$

The chemical structures represented are both C-C(= N)-H, which are equivalent. A duplication removal routine, therefore, was written to remove the different representations of the same fragment. Only one representation is output for each fragment. For fragments with a chiral centre, only one representation is output, and the two enantiomers are treated as the same.

The duplication removal subroutine has two screens. The first screen checks that the two fragments examined have the same number of the same type of atoms. If this is not the case, the subroutine jumps out of the loop and goes on to compare two other fragments. Only when the first screen gives a positive result does the subroutine compare the two fragments in detail. It permutatively rotates one fragment against the other and compares the adjacency matrices so that alternative representations can be picked out. All duplicate structures are thus located and deleted. The position-element stacks and the adjacency matrices are then re-ordered and re-numbered.

Lastly, the generated fragments are tested for unfilled valencies. If all the atoms have filled valencies (i.e., it is a molecule), then the 'fragment' is rejected. The strategy used in STRUCGEN is summarized in the pseudo-algorithm below:

(1) Choose a central atom that has not been chosen before.

(2) Add atoms to the central atom until all valencies are filled.

(3) Examine all fragments.

(a) If they have the same number of atoms then continue, otherwise GOTO 5.

(b) If they have the same numbers of different atoms then continue, otherwise GOTO 5.

(c) DO all possible permutations.
Examine if the two fragments have the same kind of bonding pattern for the same atoms. If so, this is a duplicate and only one of them should be retained. If not, continue through DO loop to try the next permutation and comparison.

(4) If 'fragment' contains unfilled valencies, GOTO 5. Otherwise discard 'fragment'.

(5) Keep fragment.

(6) Output all remaining fragments.

Ultimately 238 fragments were generated. Later it became necessary to include bivalent sulphur as one of the constituent atoms of these fragments. This was simply done by replacing the oxygen atoms by sulphur atoms one by one. 185 sulphur-containing fragments were generated, making the total number of fragments to 423.

Generation of aromatic fragments

In this work, only single-ring aromatic fragments of 5 and 6 atoms, and double-ring aromatic fragments of 9 and 10 atoms are considered. The 9-atom fragments have the atoms in a 5;6 arrangement, while the 10-atom fragments have them in a 6;6 arrangement.

In the case of single-ring and double-ring 5;6 aromatic fragments the atoms were either carbon, nitrogen or oxygen, and the generation was monitored so that the number of p electrons was $4n + 2$, where $n \notin Z^+$. This ensured that the structure was a genuine aromatic fragment. For double-ring 6;6 aromatic fragments, only carbon and nitrogen were permitted. Oxygen was not allowed because oxygen-containing 6;6 rings are not aromatic.

The aromatic fragments were generated using either pyrrole, furan, benzene, indole, benzofuran, isoindole, isobenzofuran or naphthalene as the 'parent structure'. The carbon atoms in each fragment were replaced by nitrogen one by one to give a new aromatic structure until all possible permutations had been exhausted. The number of aromatic fragments generated are listed against the 'parent structure' in Table 1. Altogether exactly 299 aromatic fragments were generated.

Generation of charged fragments

Charged fragments were not generated combinatorially because there are only a very limited number of them in existence compared to the number of neutral fragments. Any ionizable aliphat-

TABLE 1
NUMBER OF FRAGMENTS GENERATED FROM THE AROMATIC 'PARENT' STRUCTURES

Parent	Number of fragments generated
Pyrrole	10
Furan	9
Benzene	12
Indole	63
Benzofuran	63
Isoindole	36
Isobenzofuran	35
Naphthalene	71

ic fragment, from the 423 fragments, was ionized to yield charged fragments. For example, the C-NH_2 group was protonated to give C-NH_3. In addition, charged groups containing phosphorus and sulphur were included. A total of 20 charged fragments were obtained.

A total of 742 simple aliphatic, aromatic and charged fragments have been generated. Nevertheless, not all of them are frequently occurring or even chemically possible structures. We aim to include only fragments which are reasonably stable and easily synthesizable in the fragments database. These two properties are important in the manufacture of drugs. The frequency of occurrence of these fragments can reasonably be used as an indicator of their stability and perhaps ease of synthesis. The Cambridge Structural Database was used to search for the frequency of occurrence of all fragments.

Determination of fragment occurrence frequency from the CSD

In this section the QUEST programme in the CSD was used. The CONNSER commands allow one to specify a fragment by an atom properties record and by a bond properties record. The atom properties record dictates the element at each position, and the minimum number of bonded non-hydrogen atoms. The bond properties record defines the bond type between two bonded atoms, e.g. single, double, triple and delocalised bonds. A special keyword E is available to the atom properties record specification so that the minimum number of bonded non-hydrogen atoms is the exact number of bonded non-hydrogen atoms. In addition, the bonds can be specified to be all acyclic or all cyclic by the ALLBOND A or ALLBOND C keywords, respectively. Other keywords enable one to pick out fragments where all atoms in the fragment may be linked to each other only by bonds given explicitly in the bond properties records (NOLN), or where no atom in the fragment may be connected by a cyclic bond to any atom outside the fragment (NOCR).

The CONNSER command forms only part of the QUEST command related to chemical connectivities. Other parts of the QUEST command include various specifications related to the data output format, bibliographic reference, chemical and crystallographic properties of the compound of which the fragment is a part. The QUEST keywords and sub-keywords enable the user to specify, for example, the maximum atomic number of the elements present in the compound (MAXA) or the element groups present in the compound (GROUP). The keyword SCREEN also enables one to screen out certain classes of compounds e.g. SCREEN 57 would concentrate the search on organic compounds.

In our initial searches for the 423 acyclic, 299 aromatic and 20 charged fragments, the NOLN keyword was included in the CONNSER commands so that intra-fragment links were absent in the compounds retrieved. This generated the fragments having as many free valencies as possible, since the procedure maximises the number of ways the fragment can be joined to the rest of the evolving ligand. Intra-fragment links decrease the free valencies, and hence the search was limited to cases where the fragment existed without intra-fragment links. A decision about which low-frequency fragments to omit from a database of all possible fragments will always be arbitrary. Esoteric fragments of low frequency could be re-incorporated into the fragments database at a later date if this is found necessary. We decided that aliphatic fragments with less than 100 hits for the 63 000-entry Version 3.1 of CSD, should be excluded from further consideration. This cut-off point was chosen so that 99% of the total hits of all fragments could be retained.

In the case of aromatic fragments, different acceptance thresholds were applied because most aromatic rings were generally less frequently occurring than aliphatic fragments. The acceptance

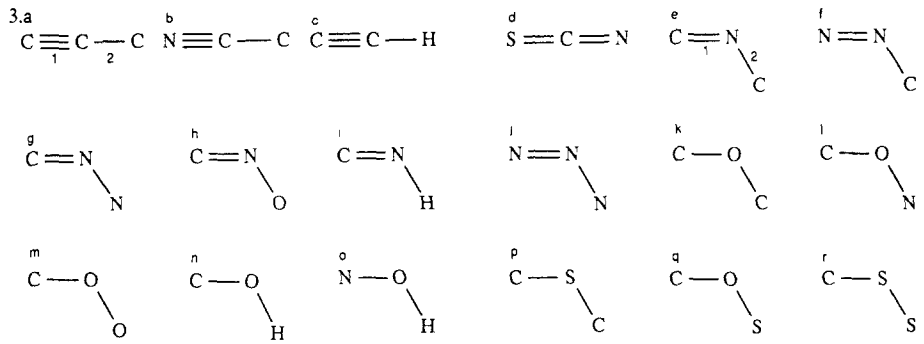


Fig. 1 The 3-atom aliphatic fragments. Fragment k represents 3 distinct fragments, depending on the hybridization state of the carbon atoms (both sp3, one sp2 and one sp3, and both sp2). Fragment n represents 2 distinct fragments, depending on the hybridization state of the carbon atom (sp2 or sp3).

threshold for single-ring 5-atom or 6-atom aromatic fragments was set at 20 hits or more, while that for double-ring 9-atom (5;6) or 10-atom (6;6) aromatic fragments was set at 15 hits or more.

In the case of charged fragments, 6 fragments achieved over 190 hits, while the rest all had less than 60 hits. The 6 fragments were accepted. Two fragments from the rejected ones, the phosphato-group and the phosphonate group, were retained specially because they were required for future work.

106 aliphatic and 8 charged fragments remained, while only 14 single-ring and 8 double-ring aromatic fragments were retrieved. A further QUEST run was executed on the 106 aliphatic fragments, this time adding keyword ALLBOND A. It was also required that only organic compounds would be retrieved (for cyclic fragments, the organic compounds screen was already set up

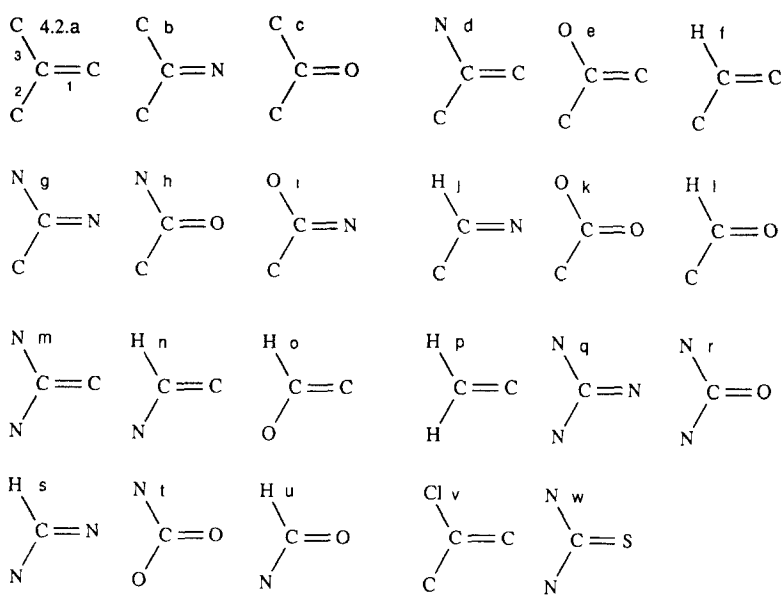


Fig. 2. The 4-atom aliphatic fragments containing one double bond

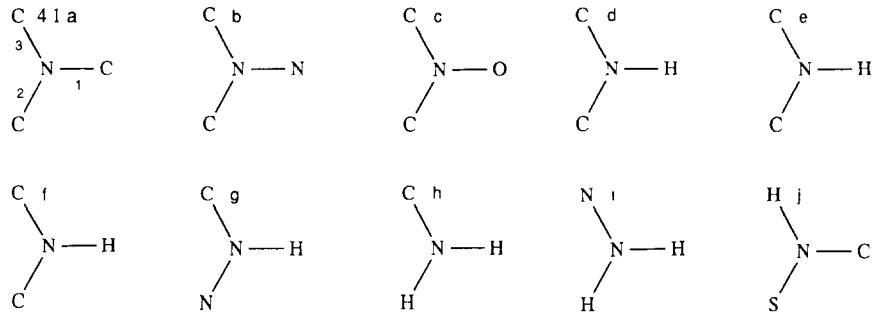


Fig. 3. The 4-atom aliphatic fragments containing 3 single bonds.

in the first round of search). After the second run, only 80 acyclic fragments were found to achieve more than 100 hits for Version 3.1 or 125 hits for Version 3.4. Subsequent fragment geometry calculations and statistics were performed only on these 80 aliphatic, 22 aromatic, and 8 charged fragments. These accepted fragments are shown in the following figures. The aliphatic fragments have been divided into 3-atom fragments (Fig. 1); 4-atom fragments containing one double bond (Fig. 2) and fragments containing three single bonds (Fig. 3); and 5-atom fragments (Fig. 4). The charged fragments are shown in Fig. 5. Five-membered aromatic rings are shown in Fig. 6, six-membered rings in Fig. 7 and fused double rings are illustrated in Fig. 8. In these figures each fragment is given a key letter and each bond is numbered in the order shown for the first fragment or explicitly for all fragments. These figures, together with their keys and bond codes, will be referred to in subsequent papers [1,2].

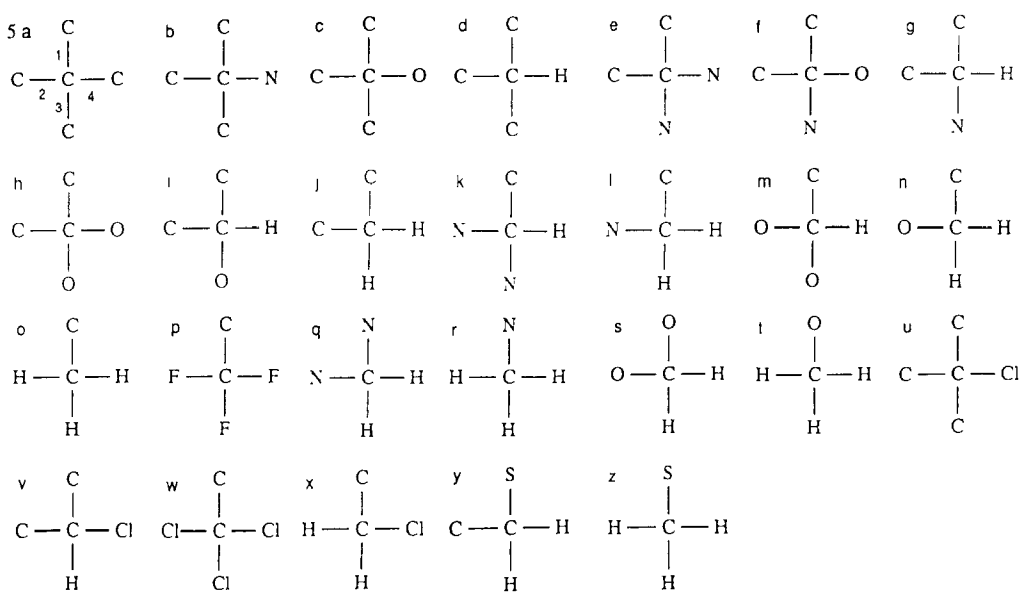


Fig. 4. The 5-atom aliphatic fragments.

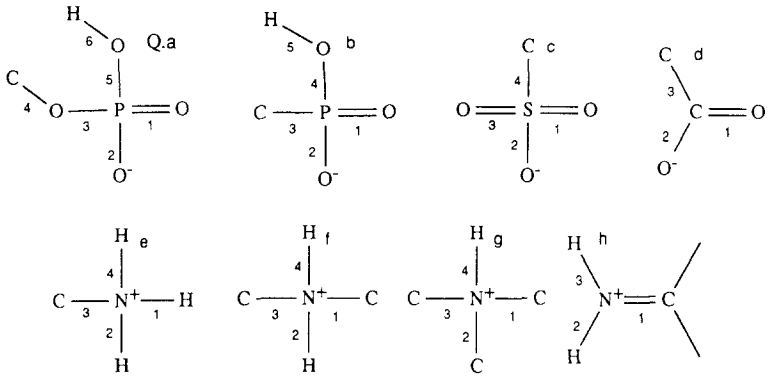


Fig 5. The charged aliphatic fragments.

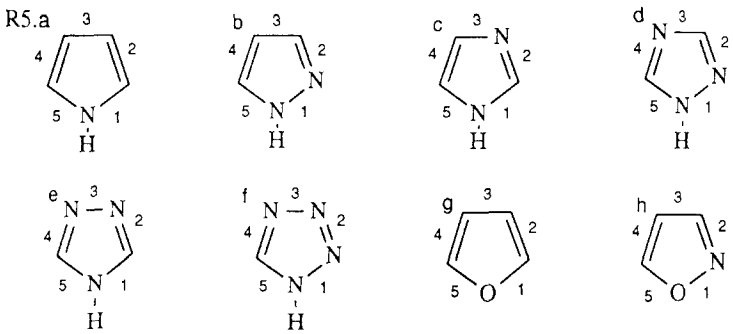


Fig. 6. The 5-membered aromatic rings.

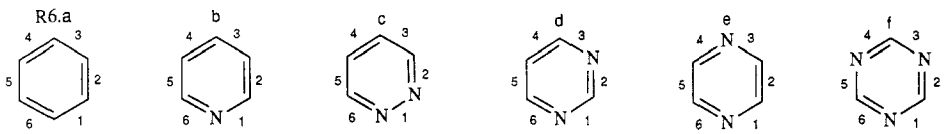


Fig. 7. The 6-membered aromatic rings.

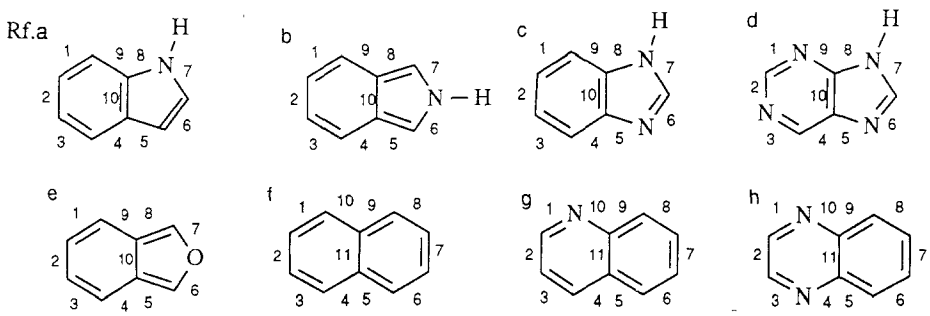


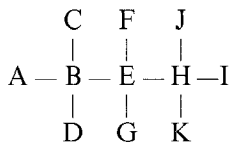
Fig 8 Fused bicyclic aromatic rings

DISCUSSION

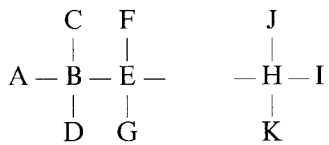
The work described in this paper has generated a large number of small fragments and has also determined the frequently-occurring ones. The size of aliphatic fragments varies from 3 heavy atoms to 5 heavy atoms, while that of aromatic fragments covers 6–10 heavy atoms. The choice of this size is determined by many factors. Ideally, we would want the fragments to be frequently occurring and thus stable. We also want this 'basic' set of fragments to be compact and yet it should be able to build up as many molecules as possible.

The larger the fragment, the more predictable are its properties. However, large fragments are usually scarcer than small fragments. If large fragments had been used, the number of fragments in the 'basic' set would have been much larger than those derived from small fragments because of the combinatoric explosion. Small fragments have the advantage that a small number of them is adequate to make up most organic molecules, and a small 'basic' set is easier to handle than a larger one. Small fragments are also more frequently-occurring.

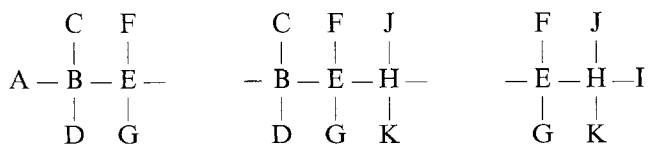
The problem of the predictability of fragment properties will be discussed in a subsequent paper [2]. Suffice it to say that even with small fragments, the properties are satisfyingly predictable. Further improvements can also be achieved by using the method of 'overlapping fragments'. For example, if one wants to build up the following structure,



then we may use two fragments, e.g.



with their individual properties to predict the properties of the whole molecule, or we could use the overlapping fragments:



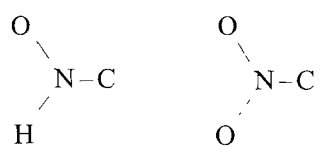
The left fragment is used to predict the properties of atoms A, B, C and D, the central fragment is used for the properties of atoms E, F and G, while the right fragment is used for atoms H, I, J and K. In this way, we would hope to achieve a higher predictability in building up a molecular structure from fragments.

In the case of aliphatic fragments, those with 3–5 atoms were chosen because they are the smal-

lest fragments with a 'central' atom to enable us to use the 'overlapping fragment' method; they are not so small as to make fragment properties unreliable in predictions. On the other hand, with aromatic fragments, each fragment is a closed aromatic system. We did try to use small basic units of 3 atoms each, but we discovered that the predictions made using these small basic units were very unreliable. For example, given the aromatic unit C-N-C, it is impossible to predict what the electronic residual charges are on each atom. We can 'close' the ring by, for example, making this 3-atom unit a part of pyridine, pyrimidine or pyrazine. However, the properties of the atoms of C-N-C change very noticeably when the fragment is in different rings. We, therefore, decided to use closed rings as our basic unit, because these are the smallest units with more constant properties. On the other hand, special fragments of this type are not so large as to be difficult to manage in a database.

Molecular fragments have a long historical usage by many research workers in different fields. Lederberg et al. [16] used them in the project DENDRAL. This project aims at writing a programme for inferring the chemical structure of a compound from its mass spectrum, and uses hierarchical logic for decision making. The various chemical fragments are analyzed by the programme, and a structure assembler then produces all possible candidates for a given mass spectrum and empirical formula. In the programme ASSEMBLE by Shelley et al. [17], fragment atoms are described using atomic descriptors, and these fragments are bonded to each other in all possible ways to see which are consistent with the chemical and spectroscopic properties obtained by experiments. The programme CHEMICS written by Sasaki et al. [18] aims at structure inference using spectroscopy data, and it has a database of 'components', partial structures of organic molecules. Lastly, the structure inference programme ACCESS of Bremser and Fachinger [19] contains a database of fragments, together with the spectroscopic properties of each fragment, for structure assembly. It can be seen that using fragments for molecular assembly has been pursued widely by different chemical disciplines to generate the fragments combinatorially and exhaustively.

Some of the fragments generated in this work have low frequencies of occurrence. There are many reasons for this. Some of them are possible geometrical fragments but chemically highly improbable. For example, the –O–O–O– fragment occurs only once in the Cambridge Structural Database. This is to be expected because it is chemically unstable. Another example are the peroxides, which are found not to be frequently occurring. In these compounds, the O–O bond is weak, so peroxides are strong oxidizing agents and are rather unstable in the presence of compounds that can be oxidized. Another group of compounds that is not frequently occurring are those containing the O–N bond, e.g.



This is again expected because nitrogen-oxygen single bonds are very weak and often decompose homolytically to form radicals.

In the following papers [1,2] the chemical fragments defined here are investigated further to describe some of their geometric and electronic properties. We would like to determine these proper-

ties for each fragment to observe how they vary from one molecule to another. This would enable us to justify whether the properties can be transferred for each fragment.

ACKNOWLEDGEMENTS

We would like to acknowledge Dr. Frank Allen of the Cambridge Structural Database for his constant patience and advice whilst this work was in progress. We are indebted to the Croucher Foundation (P.-L.C) and the Wellcome Trust for personal financial support through the Principal Research Fellowship scheme (P.M.D). Part of the work was carried out in the Cambridge Centre for Molecular Recognition supported by the SERC.

REFERENCES

1 Chau, P.-L and Dean, P.M., J. Comput.-Aided Mol. Design, 6 (1992) 397.
2 Chau, P.-L. and Dean, P M., J. Comput.-Aided Mol. Design, 6 (1992) 407.
3 Danziger, D.J and Dean, P.M., Proc. Roy. Soc. Lond., B236 (1989) 101.
4 Danziger, D.J. and Dean, P.M., Proc. Roy. Soc. Lond., B236 (1989) 115.
5 Goodford, P J., J. Med. Chem., 28 (1985) 849.
6 Boobyer, D.N.A., Goodford, P.J., McWhinnie, P.M. and Wade, R.C., J. Med. Chem., 32 (1989) 1083.
7 Lewis, R.A. and Dean, P.M., Proc. Roy. Soc. Lond., B236 (1989) 125.
8 Lewis, R.A. and Dean, P.M., Proc Roy. Soc. Lond., B236 (1989) 141.
9 DesJarlais, R.L., Sheridan, R P., Dixon, J.S , Kuntz, I.D. and Venkatargharvan, R., J. Med. Chem., 29 (1986) 2149.
10 DesJarlais, R.L., Sheridan, R.P., Seibel, G.L., Dixon, J S., Kuntz, I.D. and Venkatargharvan, R., J. Med. Chem., 31 (1988) 722.
11 Lewis, R A., J. Comput.-Aided Mol. Design, 4 (1990) 205.
12 Van Drie, J.H., Weininger, D. and Martin, Y.C., J. Comput.-Aided Mol. Design, 3 (1989) 225.
13 Rekker, R.F., The Hydrophobic Fragmental Constant, Elsevier, Amsterdam, 1977.
14 Hansch, C. and Leo, A., Substituent Constants for Correlation Analysis in Chemistry and Biology, John Wiley and Sons, New York, 1979.
15 Clementi, E., Computational Aspects of Large Chemical Systems, Springer-Verlag, Berlin, 1980.
16 Lederberg, J., Sutherland, G.L., Buchanan, B.G., Feigenbaum, E.A., Robertson, A.V., Duffield, A.M. and Djerassi, C., J. Am. Chem. Soc., 91 (1969) 2973.
17 Shelley, C.A., Hays, T.R., Munk, M.E. and Roman, R.V., Anal. Chim Acta, 103 (1978) 121.
18 Sasaki, S.-I., Abe, H., Hirota, Y., Ishida, Y., Kudo, Y., Ochiai, S., Saito, K. and Yamasaki, T., J. Chem. Inf. Comput. Sci., 18 (1978) 211.
19 Bremser, W. and Fachinger, W., Magn. Res. Chem., 23 (1985) 1056.