Structure design: An artificial intelligence-based method for the design of molecules under geometrical constraints

Alexander A. Cohen and Shimon E. Shatzmiller

Tel-Aviv University, School of Chemistry, Ramat-Aviv, Israel

This study presents an algorithm that implements artificial-intelligence[1] techniques for automated, and site-directed drug design. The aim of the method is to link two or more predetermined functional groups into a sensible molecular structure. The proposed designing process mimics the classical manual design method, in which the drug designer sits in front of the computer screen and with the aid of computer graphics attempts to design the new drug. Therefore, the key principle of the algorithm is the parameterization of some criteria that affect the decision-making process carried out by the drug designer.

This parameterization is based on the generation of weighting factors that reflect the knowledge and knowledge-based intuition of the drug designer, and thus add further rationalization to the drug design process.

The proposed algorithm has been shown to yield a large variety of different structures, of which the drug designer may choose the most sensible. Performance tests indicate that with the proper set of parameters, the method generates a new structure within a short time.

Keywords: drug design, structure generation, artificial intelligence, molecular graphs

INTRODUCTION

The design of new compounds with predetermined structures and activities has been well recognized as a key process in the development of biologically active compounds.[2] With the aid of molecular graphics,[3] the scientist is able to design a new compound simply by fitting fragments to the binding site pocket of the target molecule[4] (namely, the receptor). Even though this method is straightforward, its ability to generate feasible molecules is limited due to the difficulties in designing reasonable structures. Namely, while the assignment of key functional groups to positions on the receptor's surface is apparent (e.g., by employing a probe-surface analysis,[5] etc.), the linking of a few functional groups into a sensible molecule requires much more intensive efforts.

In the general scope of drug design, the structure-design problem could be defined as the generation of a conformationally stable molecular backbone that fits into the receptor's pocket and that positions the required functional groups in a predetermined orientation.

To accomplish this, four principal methods may be employed separately or simultaneously:

(1) *Manual drug design*. The drug designer uses graphical and geometrical tools to build a feasible structure. As stated above, the method is time consuming and rarely yields sensible structures; thus it is an unfavorable method.

(2) *Step-by-step optimization*. The structure design is carried out in a stepwise manner, where at each step the fragment that yields the minimum energy is selected from a given library of fragments.[6,7] The energy function may include nonbonded interactions[7] as well as bonded interactions.[6]

(3) *Database searching*. Geometrical and energy-based mapping of the receptor's pocket are used as input for searching substructures in a database. The best substructures found are then fused together by the drug designer to yield a molecule.[8]

(4) *Torsional-angle fitting*. This newly suggested method enables rapid conformational optimization of sequences of atoms in order to find those that comply with the specifications.[9,10] It should be noted that torsional-angle fitting is not a stand-alone method (yet), but is relevant to the structure-design problem presented above.

A precise comparison of these methods is beyond the scope of this study. However, qualitative analysis indicates

Address reprint requests to Dr. Alexander Cohen at Tel-Aviv University, School of Chemistry, Ramat-Aviv 69978, Israel.

Received 23 December 1992; revised 2 February 1993; accepted 9 February 1993

that the methods differ in the rationalization of each designing step, and in their sources of information. The distinctive differences between these methods are summarized in Table 1. Although the automatic methods (1–4) are rationalized by some sort of energy-fitting procedures, their outputs can rarely be considered to be optimal solutions. The latter statement can be justified by two reasons: One may not assume that optimization of each component yields an optimized system; and practically, it is almost impossible to take into account all the factors that may affect the structure of the new ligand.

Consequently, it has to be accepted, that the four methods generate feasible, but not necessarily optimal, solutions.

As a result, since none of these methods generates an optimal solution, and a method to deduce the optimal solution has not yet been found, one's computational efforts should concentrate on finding as many feasible solutions as possible. From the set of feasible solutions the drug designer may eliminate those that seem synthetically inapplicable, and the remaining structures can be further verified (e.g., with more comprehensive calculation methods such as molecular dynamics and, finally, synthesis and testing).

Following this analysis, a different rationalization of the drug design process should generate a different set of feasible solutions.

Hence, the manual design method, which is rationalized by intelligence-based decision-making processes, might serve as a useful source for feasible structures. Nonetheless, the main drawback of the manual design method is the time required to complete a designing session. Consequently, we have decided to develop a method that parameterizes the designer's decision-making processes, and thus enables automatization of the manual design method. The following sections focus on the principles of the method, and some test cases that illustrate its applicability to the drug design problem.

In order to automate the manual design method, it is best to disassemble it into several elementary, stand-alone components or steps. Subsequently, we have to consider how each stand-alone step involves decision making, and attempt to parameterize the decision criteria. Finally, the components should be reassembled together to yield a complete method.

The manual-design method could be simplified and disassembled into components as follows:

(1) Assignment of the functional groups. This step is considered to be the most important task during the drug design, since it determines the functioning of the new drug. Furthermore, the orientations of these groups mark out the geometrical constraints that should guide the design of the bridging part. However, as mentioned earlier, this assignment can be carried out with the aid of automatic well-known algorithms. Moreover, if an X-ray structure of the receptor-linked complex is available, the drug designer usually uses, for this purpose, the same binding sites used by the native ligand.

(2) Selecting one group as the starting point. After positioning the functional groups, the designer starts to link them together. To perform this linking, the designer decides arbitrarily where he wants to start.

(3) Adding a single atom or a small fragment. During this stage, the designer selects an atom or a small fragment and fuses it to the open valence of the preselected first functional group. Although this step might seem rather trivial, it is based on the designer's best chemical knowledge and intuition.

(4) Add/Delete sequence. The drug designer repeats step 3 a few times and then he decides whether the growing chain looks sensible. If not, a few atoms are deleted and a few others are added, etc. This stage is usually the "exhausting" part of the process.

(5) Reaching the next functional group. After many consecutive add/delete sequences, the open side of the growing chain finally reaches the next functional group. The drug designer closes the open valences, and a molecule is formed.

(6) General overview. The drug designer verifies the proposed structure, and makes some corrections here and there. These corrections include mainly conformational locking (e.g., the insertion of a "benzo" moiety in order to preserve the conformation of conjugated double bonds).

(7) Optimization. The new molecule that appears on the screen can be optimized with some sort of optimization method, for example molecular mechanics.[11] The drug designer considers the optimized geometry, the resulting heat of formation of the new molecule, and the binding energy, in order to decide whether to accept the new structure, or to start the process all over again.

According to this scheme, artificial intelligence could be useful for the automation of steps 3–5. It should be noted that steps 1 and 2 might as well be implemented by an artificial-intelligence scheme. However, the assignment of the key groups (step 1) is, as stated above, best achieved by

Table 1. Rationalization and source of information of the four basic structure-design methods

		Structure design methods		
Criterion	Manual design	Step-by-step optimization	Database searching	Torsional-angle fitting
Rationalization	Human intelligence	Molecular mechanics	Energy fitting	Energy fitting
Source of information	Unlimited	Library of fragments	X-ray data[a]	Atom palletes
Designing speed	Very slow	Medium	Fast	Very fast

[a] From the Cambridge database, or any other database of three-dimensional structures.

numerical methods, while the selection of the starting point (step 2) is arbitrary and thus might as well be performed interactively. The next sections describe in detail the approach used to automate steps 3–5.

METHODS

Automated decision-making process

Definitions Let us define three types of decision-making processes (DMPs) that are relevant to the structure-design problem:[12]

- Λ_a. (Literally, a decision making process of type *A*). A boolean DMP that is based on the state of each relevant criterion that may affect it. The combined state of all criteria determines the decision.
- Λ_{aa}. This DMP is much like Λ_a, but the corresponding criteria only make a preference for one decision, rather than dictating it.
- Λ_b. Represents a selection-based DMP. Namely, the decision is concerned with selecting one option among several possible options.

Mathematical formalism Λ_a and Λ_{aa} could be translated into a formal mathematical syntax in two steps:

(1) Consider each factor that may affect the decision, and assign to it a *weight*. The weight is a numerical parameter that reflects the influence of the corresponding factor on the final decision, regarding the whole problem.
(2) Combine the weights to produce a score (the combination is not necessarily linear). If the weights and the corresponding combination are normalized, then the score ranges from 0 to 1.

Let Q be an event that depends on a set of factors $\{F\}$, where $\{W\}$ is the corresponding set of weights. The score is given by a combination of $\{W\}$, namely $g\{W\}$, where $g\{W\} \in [0,1]$. Λ_a (e.g., accept or reject Q) is carried out by setting a threshold value T as follows:

$$Q = g\{W\} \geq T \,?\, \text{accept} : \text{reject} \tag{1}$$

Please note that the structure of Equation (1) is based on C syntax; i.e., *value = condition ? value if true : value if false*.

Following Equation (1), a decision based on Λ_{aa} could be simulated by replacing T with a random number R. Thus, if R is a random number equally distributed on [0,1] then Equation (1) may be rewritten to yield:

$$Q = g\{W\} \geq R \,?\, \text{accept} : \text{reject} \tag{2}$$

The formalism of Λ_b might be as follows: Let F_i represent a possible selection for Q, and W_i the normalized probability that F_i will be the selection, namely $\Sigma W_i = 1$. Let R be a random number equally distributed on [0,1]. Then

$$Q = R \in [S_{i-1}, S_{i-1} + W_i) \,?\, F_i : \text{i++} \tag{3}$$

where S_{i-1} stands for the sum of the first $(i - 1)$ weights. The symbol $i++$ is literally compatible with "increment i by 1." Thus, Equation (3) represents a loop over all possible $\{W\}$ until a match is found. If the number of possibilities equals 2, then Λ_b reduces to Λ_{aa}, Equation (2). Please note that even though $\{W\}$ might not be sorted, proper use of Equation (3) requires consistent order of access.

Implementation to the drug design problem

Selecting the next atom or fragment The selection of the next atom or fragment corresponds to Λ_b. For that case, a look-up table is built (prior to the design), and all the possible bonds that could be formed between the atom at the chain terminus and any other atom are ranked according to the probability of bond formation. By applying Equation (3), the resulting output is the type of the new bond. Table 2 illustrates a model of a bonding probability table (BPT). The assignment of the bonding probabilities is based on our intuition, and clearly, others might think of different values. With the set of the bonding probabilities given by Table 2, the probability of creating simple linear hydrocarbons is demonstrated by Table 3.

After the next atom has been selected, its orientation must be specified. Bond lengths and bond angles could be assigned directly from tables containing internal relationships data. (In this work we used CHARMm™ parameters.[11]) The dihedral angle is assumed to be a minimum, but not necessarily the global minimum. For example, the dihedral angle between atom 1 and 4 in a linear molecule 1-2-3-4, where atoms 2 and 3 are in SP^3 hybridization might be 60°, 180° or 300°. Λ_b, with equal probabilities, is used to assign atom number 4 to its position. Thus, while the atomic sequence is determined by the BPT, the conformation is depicted randomly. It should be noted that the configuration and the conformation construction is approximated by the rigid geometry approximation. This approximation was recently validated by Palmer and Scheraga.[9]

Verification of the selection The second decision that we have to make is whether to accept or reject the new designed fragment. This decision is concerned with the geometric constraints imposed by the shape of the receptor, the steric strain, and the desired goal. Consequently, we may present the constraints by assigning four criteria that enable correct decision making.

The first criterion checks whether the growing chain grows toward the target functional group. The corresponding weight must take into account the relative distances of all other atoms to the target. This criterion could be parameterized, for example, by Equation (4):

$$W_r = [N_d/(N_c + 1)]/N = (N - N_c)/(N_c + 1) \tag{4}$$

Table 2. Model of a bonding probability table. The first row indicates the current atom in the chain terminus, and the first column represents the suggested new atom

	C-sp^3	[a]C-sp^2	[b]C-sp^2	[c]C-sp^2
C-sp^3	0.8	—	0.6	—
[a]C-sp^2	0.2	—	—	—
[b]C-sp^2	—	1	—	1
[c]C-sp^2	—	—	0.4	—

[a] The first sp^2 atom in a double bond.
[b] The second sp^2 atom in a double bond.
[c] The third sp^2 atom in a conjugated diene.

Table 3. Probability to design simple, linear hydrocarbons that contain a specified number of carbon atoms, by employing the bonding probabilities listed in Table 2. The first atom in each molecule was assigned to be a sp^3 carbon atom

Name[a]	Number of designed atoms	Probability
Butane	3	0.512
2-Butene	3	0.120
3-Butene	3	0.160
Pentane	4	0.410
2-Pentene	4	0.096
3-Pentene	4	0.096
4-Pentene	4	0.128
2,4-Pentadiene	4	0.080
Hexane	5	0.328
2-Hexene	5	0.077
2,5-Hexadiene	5	0.024
2,4-Hexadiene	5	0.048
Heptane	6	0.262
2-Heptene	6	0.061
2,5-Heptadiene	6	0.019
2,6-Heptadiene	6	0.019
2,4-Heptadiene	6	0.038
2,4,6-Heptatriene	6	0.032
Octane	7	0.210
2,4,7-Octatriene	7	0.029

[a] Numbered always from the first c-sp^3 atom.

where W_r is the weighting factor (the subscript r stands for *relative*); N_d is the number of atoms in the chain that are further distant than the Nth atom; N_c is the corresponding number of atoms that are less distant; and N is the total number of atoms in the growing chain. Clearly, Equation (4) yields unity (the maximum) if N_c equals zero, and zero if N_d is zero.

The second criterion monitors the real distance to the target. This factor increases the probability that a new atom positioned near the target is going to be accepted. For that case we have defined a distance threshold D_t to yield the corresponding weighting factor,

$$W_a = d > D_t \,?\, d/D_t : 1.0 \tag{5}$$

where W_a is the weighting corresponding to the absolute distance; d is the distance from the target; and D_t is the threshold distance. The value of D_t is indicative of the tolerance level of the design and thus could be used as the "successful termination" criterion. The value of W_a could be useful for determining the completion of the process.

The designing process should preclude collisions between the new chain terminus and other parts of the chain. Consequently, the third criterion was implemented by:

$$W_c = \text{New bond intersects with the chain} \,?\, 0 : 1 \tag{6}$$

The term *intersects* refers to the van der Waals interaction, as approximated by the hard-spheres assumption, between the chain's terminus and any other part of the chain.

Finally, the fourth criterion considers the interaction between the chain and the host receptor. In other words, the conformation of the new terminus should obey the shape structure of the host. This criterion is given by Equation (7):

$$W_p = \text{New bond penetrates the receptor's domain} \,?\, 0 : 1 \tag{7}$$

The four criteria presented by Equations (4–7) are encapsulated by the normalization function g to yield the corresponding score:

$$g(W_r, W_a, W_c, W_p) = W_c W_p (W_a + W_r)/2 \tag{8}$$

Where g is implemented by a Λ_{aa} scheme. Note, however, that W_c and W_p are implicitly related to Λ_a, since their values are either zero or unity.

Recursive corrections Early experiments showed that by using Equation (8) alone as the scoring basis for the designing process the method rarely converged to a molecule. Further examination indicated that since the probability of making a "bad move" is not zero, in most cases the new chain grew toward the wrong direction.

Consequently, we have decided to implement a recursive correction mechanism. The procedure used applies the Λ_{aa} scheme and a parameter called the *absolute rejection probability*, P_r, in order to decide whether to reject or accept the last atom assigned to the chain. The decision making process was carried out in every cycle.

Unfortunately, the value of P_r must be deduced experimentally. However, as illustrated by Test 2 below, the range of possible values is relatively wide.

The proposed algorithm The proposed algorithm assumes that the drug designer has already assigned the key groups to their positions in the binding site. The next step is interactive, where the drug designer selects the exact starting and ending loci. The relative orientation of these points is determined by the nature of the corresponding functional group and is translated into Cartesian coordinates. These points are used as the chain edges. Afterwards the automated part of the manual design is carried out by implementing sequentially a group selection, verification, and recursive correction. Finally, the drug designer regains control and is able to interactively verify the proposed structure. Figure 1 illustrates the algorithm derived from this scheme. As stated previously, there might be an infinite number of ways to define the weights and to combine them into normalized forms. However, the proposed algorithm neither depends on the form of the weights, nor on the form of the scoring system, as long as they are normalized and well behaved.

RESULTS

All calculations were performed on a SGI-4D™ (Silicon Graphics™ Inc.) computer, under the UNIX™ operating system. Internal relationships (i.e., bond length, bond angles and dihedral angles) and nonbonding parameters were implemented from CHARMm™ standard parameters.[11] The random number generator used was the standard generator attached to the C compiler of the operating system. For the following feasibility study, several simplifications were employed: The interaction with the receptor has been neglected

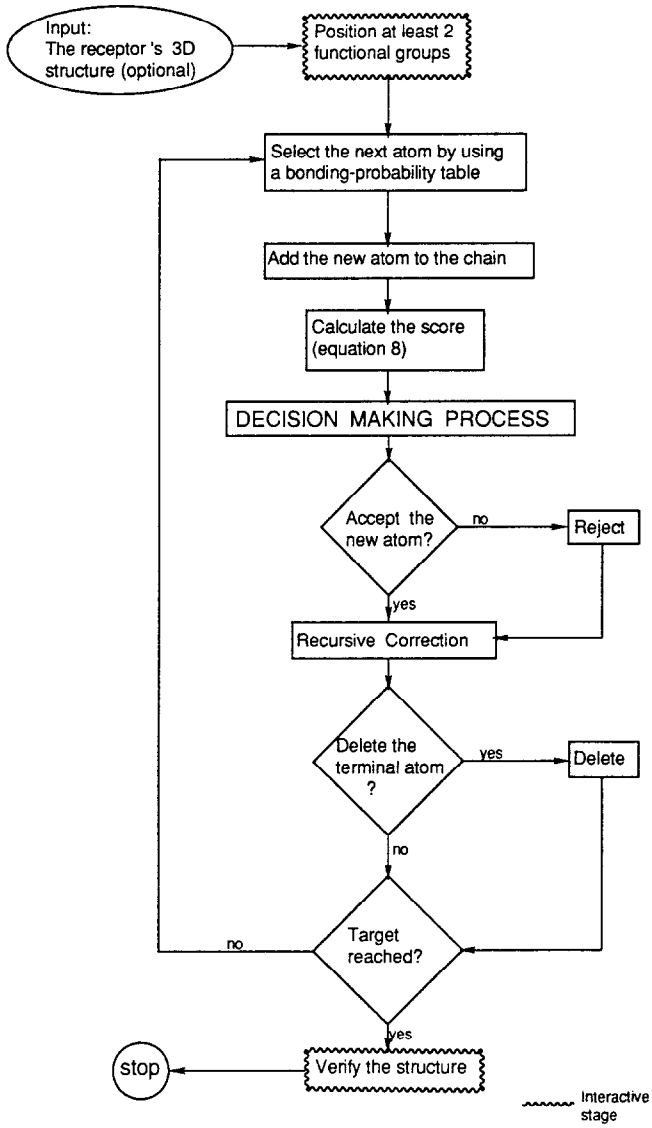

Figure 1. Proposed structure-design algorithm.

($W_p \equiv 1$); the program used only H, C(sp^3) and C(sp^2) atoms as the building blocks; and the tolerance D_t was set equal to 1.5 Å, i.e, approximately a C-C bond length.

The aim of the design was to find a structure that could bridge over two functional groups separated by 7.8 Å (chosen arbitrarily). The process assumed to diverge if the number of atoms in a chain exceeded 50, or the number of attempts exceeded 50,000.

Test 1. Building molecules

Our first test case considers the design of a chain of atoms composed only of sp^3 and sp^2 carbon atoms and additional hydrogens. The BPT used is as illustrated by Table 2, but the probability of producing a conjugated diene was reduced from 0.4/0.6 to 0.1/0.9. Due to the random nature of the designing scheme the seed number that initializes the random number generator was arbitrarily set to four different seeds. The results are shown in Figure 2.

Test 2. The role of the recursive correction

The second example illustrates the role of the recursive correction scheme. The absolute probability of any chain terminus that was previously accepted was varied from 0.05 probability (to be rejected) to 0.85. Table 4 compares the number of atoms in the designed structure and the number of attempts required, as a function of P_r.

Test 3. Manipulating the selection criteria

While the previous examples examined the numerical behavior of the algorithm, this test inspected the role of the BPT. Manipulation of this table enables the drug designer to direct the frequency of atoms' types in the designed molecule.

In the previous tests, the probability that an sp^2 atom be added to the chain was 20%. Figure 3 illustrates the structural changes followed by increasing this probability to 40%, 60% and 80%. The other run-time conditions remained as in Test 1.

DISCUSSION

In this study we have described a method that enables the design of molecular structures under constraints. The illustrative examples given above indicate that, in principle, the method is capable of designing novel structures that satisfy a set of given restrictions.

The results of the Test 1 represent the random nature of the proposed algorithm. Any change of the random generator seed number may lead into a different structure. Even though systematic methods are generally preferable over random-based methods, it is argued that a systematic method may not lead necessarily to an optimal solution. Furthermore, the known systematic methods which use databases, libraries of fragments, etc., are by default limited by their sources of information. By contrast, a random-based method yields, inherently, unbiased structures. Moreover, the irregularity of the proposed structure might even refresh the "reservoir" of feasible structures.

A second consideration is the ability of the drug designer to influence the proposed structure by manipulating the BPT. This feature of the artificial intelligence-based method does not contradict the aforementioned inherent novelty of the proposed structure, since the drug designer may only affect the composition, but not the length, the three-dimensional structure, etc. As illustrated by Figures 2 and 3, the variety of the designed structures is mainly controlled by the parameters used, i.e., the drug designer's will.

It is generally expected that as the constraints become more strict, the number of feasible structures shall be reduced. For example, an explicit consideration of the receptor might enforce a more direct approach of the chain toward the target and thus avoid too peculiar structures, such as structure III in Figure 2. As a result, the number of random-search cycles might increase. Nonetheless, since the number of different combinations of atoms' sequences is infinite, the rationalized reduction of possibilities remains the primary advantage of the method.

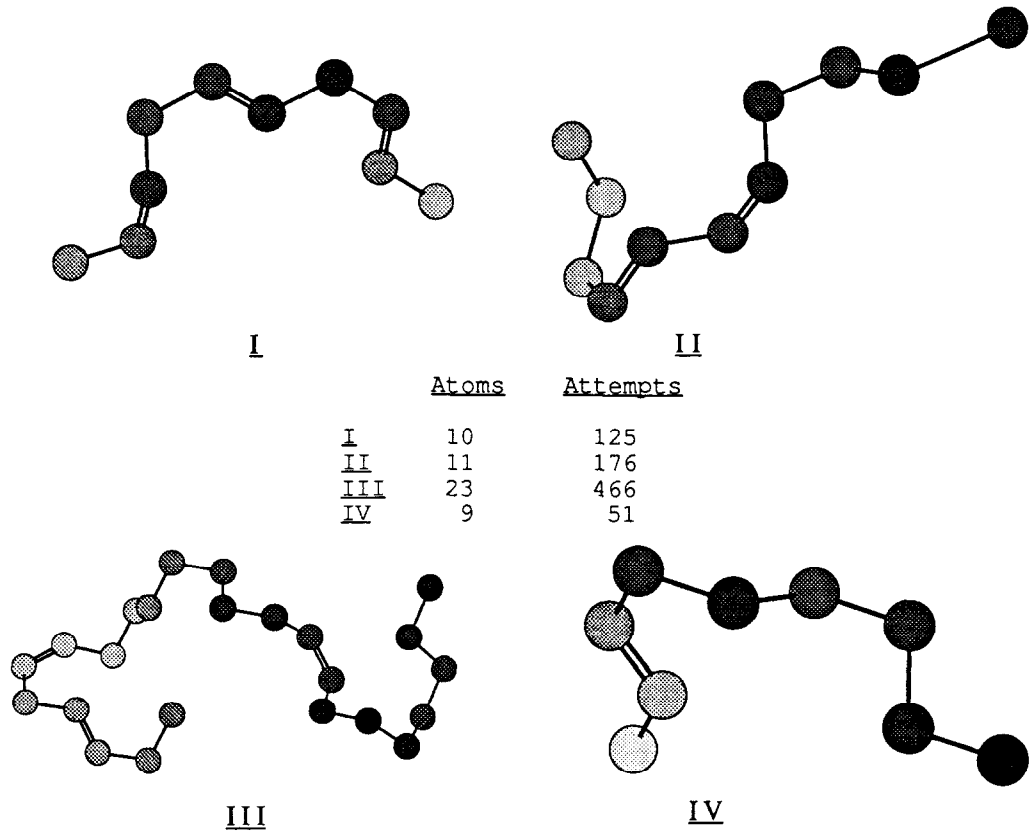

Figure 2. Design of four different structures with four different seed numbers. The atoms are shaded by their depth (dark circles in front)

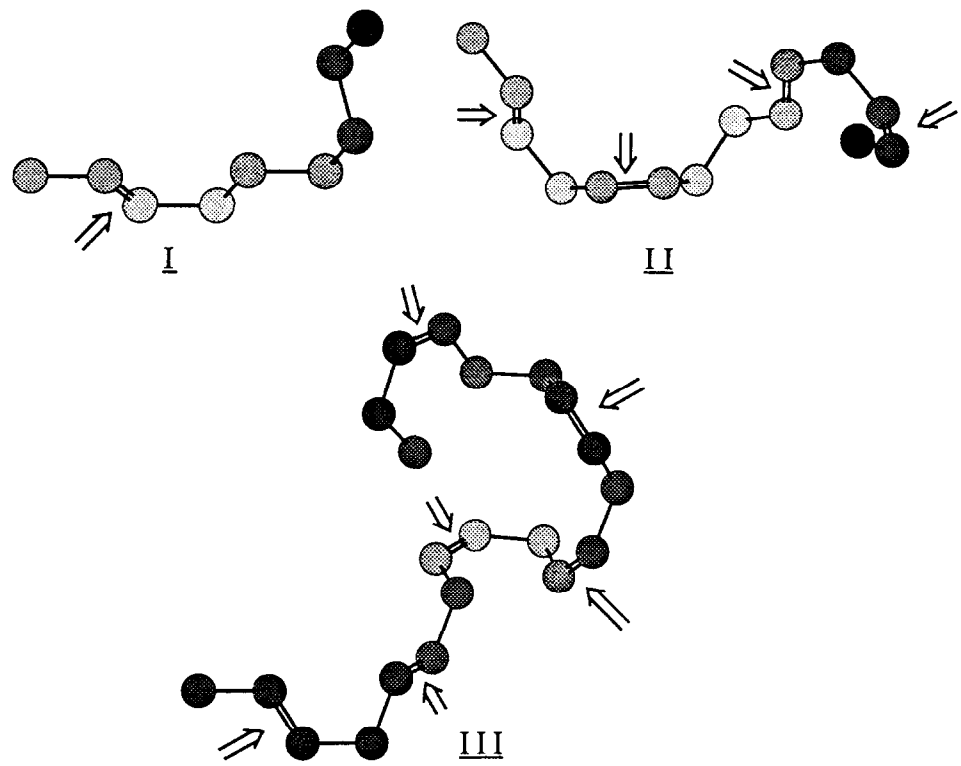

Figure 3. Design of three different structures with increasing probability of double bonds (not conjugated): (I) 40%, (II) 60%, (III) 80%. The resulted sp^2 frequency is: (I) 22%, (II) 57%, (III) 57%. Arrows point to the double bonds. The atoms are shaded by their depth (dark circles in front)

The bond lengths, bond angles and dihedral angles are assigned with their natural values. Clearly, the dihedral angle is the principal parameter that governs the three-dimensional structure of the new structure. Therefore, the method distinguishes between all possible (native) dihedral angles. Thus the order in which atoms are added to the open valences of the atom in the chain's terminus imposes prochirality, cis-trans differentiation, etc. It should be noted that by using a random-based conformational design there is a nonzero probability of designing conformationally unstable structures (e.g., with too many gauche interactions). One way to reduce the probability to obtain such structures is to include a penalty parameter that takes into account the torsional energy. However, since the conformation of any linear chain is *a priori* ambiguous (i.e., due to dynamical changes), conformational locking is an essential post-process. Therefore, design based on torsional energy seems unneeded.

Numerical stability

The structure of the proposed algorithm raises the problem of numerical stability, i.e., the convergence of the process into a satisfying structure. Since the method depends on many parameters, some of which may not be optimized, in some cases the method won't converge. Figure 4 illustrates a representative convergence scheme of one of the previous tests. The "erratic" behavior of the graph in Figure 4, implies that since a convergence was reached by chance, the method might as well diverge. However, although the convergence problem is intrinsic, the drug designer may alter one or two parameters in order to attain convergence. The results shown in Table 4 indicate how the process convergence is affected by the value of the absolute rejection probability. Manipulation of the BPT, or even alteration of the random number generator's seed number, is expected to cause significant changes to the numbers in Table 4. However, a value of 0.25–0.65 seems to be a good initial guess for the absolute rejection probability.

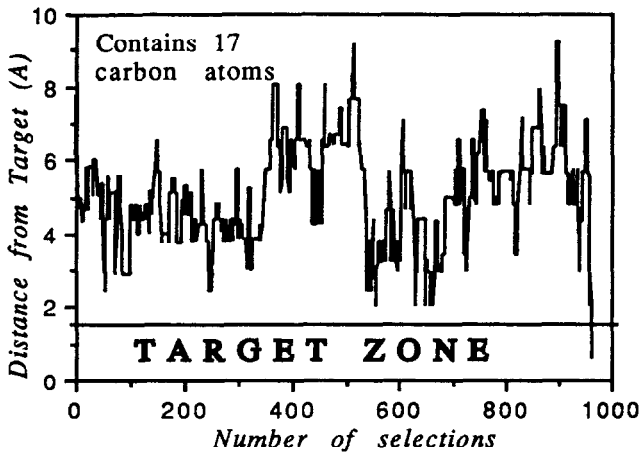

Figure 4. Typical convergence scheme. The graph shows the propagation of the chain's growing terminus toward the target group. The area marked Target zone *indicates a distance of 1.5 Å or less from the position of the target group*

Table 4. Role of the recursive correction scheme

Absolute rejection probability	Number of atoms in the chain	Number of attempts
0.05	Not converged	(Too many atoms)
0.10	47	3241
0.15	28	1280
0.20	14	297
0.25	9	87
0.30	10	125
0.35	8	226
0.40	13	276
0.45	10	197
0.50	9	176
0.55	8	28
0.60	11	110
0.65	14	156
0.70	8	2385
0.75	8	30500
0.80	Not converged	(Too many cycles)

CPU considerations

If the input constraints allow many degrees of freedom for the designing process, some structures might look quite peculiar, and practically, inapplicable. Consequently, the time required to generate a feasible structure is very important, since it determines the number of feasible structures that the drug designer may generate during a reasonable period of time. With the simplifications stated above, the CPU time required to perform 20,000 consecutive attempts to add an atom to the growing chain was approximately 1.0 minute (tested on the IRIS 4D™ computer). It should be noted, however, that explicit consideration of the binding energy might increase significantly the time required per attempt. As illustrated by the number of attempts per design represented in Table 4, the process converges within seconds.

CONCLUSION

In this communication we have presented an artificial intelligence-based method for the design of structures under geometrical constraints. The proposed algorithm is based on the translation of the drug designer's intuition and the constraints into a computerized syntax. With the proper choice of the weights, the normalization functions, and the flow-control parameters, the automatic process creates feasible structures that comply with the specifications. Moreover, the manipulation of these variables enables a large variety of rationalized structure models.

The inherent flexibility of the proposed method enables the insertion of a few other constraints, e.g., explicit consideration of steric interaction, explicit determination of binding energy, etc.

Future developments of the method requires rigorous treatment of the parameterization following by "field tests," i.e., *de novo* design of real drugs.

In summary, it has been demonstrated that a non-systematic intuition-oriented, but yet simple, conformational search could be useful for rapid structure-design applications.

REFERENCES

1 In the scope of the following study, the term *artificial intelligence* is used to describe the parameterization and subsequent computation of the decision-making processes involved in drug design.

2 For a pioneering research in the field see: Blaney, J.M., Jorgencen, E.C., Connoly, M.L., Ferrin, T.E., Langridge, R., Oately, S.J., Burridge, J.M., and Blake, C.C.F. Computer graphics in drug design: Molecular modeling of Thyroid hormone—Prealbumin interactions. *J. Med. Chem.* 1982, **25,** 785–790

3 For example, two commercially available softwares: QUANTA™, Polygen™ Corporation, 200 Fifth Avenue, Waltham, MA 02254, USA; and INSIGHT II™, BIOSYM Technologies, Inc., 9685 Scranton Road, San Diego, CA 92121, USA

4 Kollman, P. Theory of complex molecular interactions: Computer graphics, distance geometry, molecular mechanics and quantum mechanics. *Acc. Chem. Res.* 1985, **18,** 105–111

5 Goodford, P.J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 1985, **28,** 849–857

6 Crippen, G.M., and Havel, T.F. *Distance Geometry and Molecular Conformation.* Research Studies Press LTD, John Wiley & Sons Inc., New York (1988), Chap. 8, pp. 361–429

7 Bohm, H.J. The computer program Ludi: a new method for the *de novo* design of enzyme inhibitors. *J. Comp. Aided Molec. Design,* 1992, **6,** 61–78

8 Meny, E.C., Shoichet, B.K., and Kuntz, I.D. Automated docking with grid-based energy evaluation. *J. Comp. Chem.* 1992, **13,** 505–524; Shoichet, B.K., Bodian, D.L., and Kuntz, I.D. Molecular docking using shape descriptors. *J. Comp. Chem.* 1992, **13,** 380–397; Leach, A.R. and Kuntz, I.D. Conformational analysis of flexible ligands in macromolecular receptor sites. *J. Comp. Chem.* 1992, **13,** 730–748

9 Palmer, K.A. and Scheraga, H.A. Standard-geometry chains fitted to X-ray derived structures: Validation of the rigid-geometry approximation. I. Chain closure through a limited search of "loop" conformations. *J. Comp. Chem.* 1991, **12,** 505–526

10 Lewis, R.A. Automated site-directed drug design: a method for the generation of general three-dimensional molecular graphs. *J. Mol. Graphics.* 1992, **10,** 131–143

11 CHARMm™, Polygen™ Corporation, 200 Fifth Avenue, Waltham, MA 02254, USA. For the mathematical details: Brooks III, C.L., Karplus, M., and Pettitt, B.M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. Advances in Chemical Physics.* Volume LXXI. John Wiley & Sons, New York (1988)

12 The mathematical formalism used to describe the decision-making processes is, as far as we know, applied for the first time for this kind of decision. However, analysis of decision-making processes is a rather basic concept in the field of artificial intelligence.