Starting Point to Molecular Design: Efficient Automated 3D Model Builder Key3D

Miho Yamada MIZUTANI, Kensuke NAKAMURA,[1] Tazuko ICHINOSE, and Akiko ITAI*

Institute of Medicinal Molecular Design, Inc.; 5–24–5 Hongo, Bunkyo-ku, Tokyo 113–0033, Japan.

Received August 4, 2006; accepted October 2, 2006; published online October 6, 2006

Obtaining three-dimensional (3D) structures from structural formulae is a crucial process in molecular design. We have developed a new 3D model builder, Key3D, in which the simplified distance geometry technique and structure optimization based on the MMFF force field are combined. In an evaluation study using 598 crystal structures, the high performance and accuracy of Key3D were demonstrated. In the "flexible-fitting" test, which is focused on practical usefulness in the molecular design process, 88% of the Key3D structures acceptably reproduced the reference crystal structures (root-mean-square deviation $<$0.6 Å) upon rotation of acyclic bonds. These results indicate that Key3D will be very effective in providing starting points for practical molecular design.

Key words three-dimensional (3D) structure; drug design; Key3D; distance geometry; MMFF force field; MO calculation

Three-dimensional (3D) molecular structures are required in many fields of molecular design, such as docking simulation, virtual screening, molecular superposition and evaluation of quantitative structure–activity relationships, and methods to generate 3D structures of small molecules from two-dimensional (2D) structures play a key role in providing starting points.

A number of automated 3D model builders have been developed.[2] CORINA,[3] CONCORD,[4,5] and CONVERTER[6] are the most widely used methods. Two representative approaches to create 3D structures are a rule-based algorithm and the use of the distance geometry method. The rule-based approach is very rapid in general, and high accuracy can be expected with molecules for which all the components are incorporated in the internal rules and library. However, this approach is inevitably restricted by the number and quality of internal rules and the library size. Methods based on a distance geometry algorithm can handle a wide range of structures without the limitations of the rule-based approach, but require longer computational time.

Although various 3D model builders have been developed and subsequently improved in various respects, these methods may still not necessarily be adequate for providing starting points for actual molecular design. For instance, it has been reported that the accuracies of docking models which are generated starting from the CORINA structures are lower than those obtained using the structures optimized with the Merck molecular force field (MMFF94),[7—11] even though CORINA is one of the most sophisticated methods currently available.[12] We set out to develop a new 3D model builder with particular emphasis on the generation of high-quality 3D structures that would be suitable for molecular design. We named our method Key3D.

In developing Key3D, we assigned high priority to both accuracy of the converted 3D structure and the conversion rate. We chose an approach based on distance geometry technique, because it does not depend on a fragment library from which data may be missing. Regarding accuracy, we placed stronger emphasis on the accuracy of bond lengths, bond angles and ring conformations than on the conformations of acyclic moieties. The acyclic portion of ligand molecule has large conformational flexibility, and the conformation in the free state is often very different from that in the bound form to the target protein. So, the bound conformation of the ligand should be predicted in the molecular design stage, *e.g.*, by docking or molecular superposition, taking into account the intermolecular interaction with the target protein. Currently, many programs for molecular design can deal with conformational flexibility of an acyclic moiety of a ligand.[12—17] Besides structural accuracy and conversion rate, computational speed is also an important factor. However, there is a trade-off between structural accuracy and calculation speed, and in Key3D, structural accuracy is given priority over speed. This is because the generation of 3D structures needs to be performed only once when constructing a 3D database, and the 3D structures stored in the database can be used as many times as needed. Nevertheless, we tried to speed up the calculation as much as possible by significant simplification of the distance geometry procedure. Also, to enhance the quality of the created 3D structures, structure optimization is performed with MMFF94, which offers high accuracy and reliability in handling small organic molecules. In addition, molecular design based on 3D structures requires the proper handling of stereochemistry (*i.e.*, generation of structures with correct stereochemistry and enumeration of possible stereoisomers) and control of the protonation state, depending on the calculation model. Key3D was designed to meet these requirements.

In this paper, we describe in detail the Key3D algorithm. Then, we evaluate the performance of Key3D. For assessing the validity of the Key3D structures as inputs to molecular design tools, we have developed a "flexible-fitting test" method where the accuracy in terms of only bond lengths, bond angles and ring conformation can be estimated by allowing free rotation of acyclic bonds. Use of the Key3D structures as inputs for ADAM flexible docking, and validation of the docking accuracy have already been reported.[18]

Methods

Flowchart of Key3D Figure 1 shows a flowchart of the Key3D method. Letters in parentheses in Fig. 1 correspond to the following sections.

(a) The structural formula prepared by using chemical structure drawing tools such as ISIS/DRAW (MDL Information Systems, Inc.) or ChemDraw

---

* To whom correspondence should be addressed. e-mail: itai@immd.co.jp

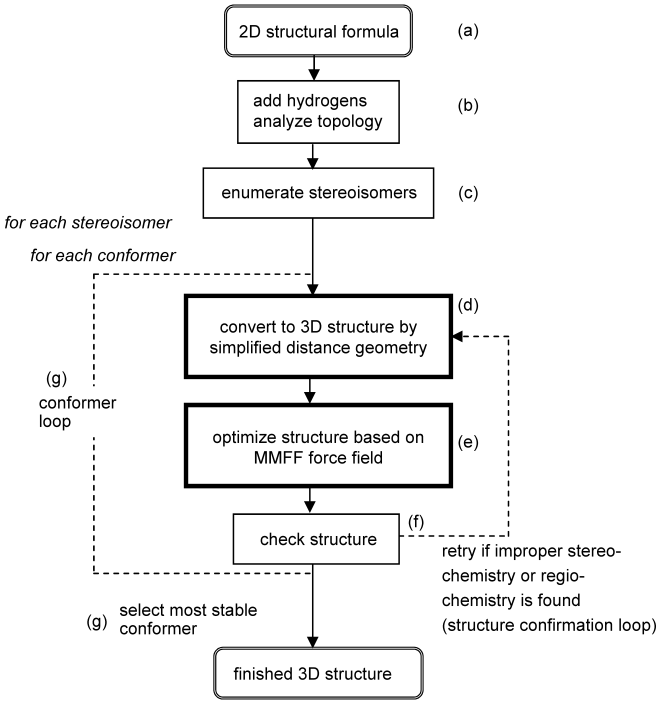

Fig. 1. Flowchart of Our 3D Model Builder, Key3D

Letters in parentheses correspond to the descriptions in the text.

(CambridgeSoft.Com) should be stored in the appropriate file format, *i.e.*, MDL format or MOL2 format of Tripos. Key3D reads this 2D topology data. Only atomic species and a connection table for the non-hydrogen atoms are needed as the 2D information. The elements that Key3D can deal with are H, C, N, O, Si, P, S, F, Cl, Br and I. Other elements are rarely found in general drug-like and bioactive molecules.

(b) The ring systems are identified, and the functional groups for which neutral and ionized forms are possible, *e.g.*, carboxylate, phosphate and amidino groups, are extracted. The protonation states for these groups can be designated by the user. According to the bond order and the designated protonation states, hydrogen atoms are added to the input structure.

(c) Analysis of stereochemistry is performed, and the stereoisomers to be generated are enumerated. When the stereochemistry is clearly defined in the input data (*e.g.*, as the bond records), Key3D creates the 3D structure of the corresponding isomer. If the molecule contains asymmetric centers with ambiguous stereochemistry, all the possible stereoisomers can be generated. However, when a large database has to be handled, it is impractical in terms of disk space to generate all possible 2^n isomers for n asymmetric centers of ambiguous configuration. So, the number of chiral centers to be considered in creating stereoisomers is limited to a maximum of six in the current version, and the user also can control the number of isomers to be generated.

(d) For each stereoisomer, a 3D structure is constructed by means of the simplified distance geometry method, as described later.

(e) The constructed 3D structure is subjected to structure optimization based on MMFF94, to improve its quality. The original parameter set of MMFF94 accommodates a wide variety of small organic molecules, but not all. We have added some parameters to the original set so that most molecules in commercial compound libraries, such as Available Chemicals Directory (MDL Information Systems, Inc.) and MAYBRIDGE database, can be handled properly.

(f) Key3D examines whether the stereochemistry and regiochemistry of the generated 3D structure are consistent with the expected ones. If any inconsistency is found, Key3D retries the construction of 3D structure by returning to the distance geometry step (this cycle is called the "structure confirmation loop").

(g) The cycles of generation of 3D structure by the distance geometry method and structure optimization (d—f) are iterated several times, and as the result, a given number of conformers is produced for each stereoisomer ("conformer loop"). In general, satisfactory results can be obtained through three to five iterations. The conformer with the lowest MMFF energy is selected for each stereoisomer and output. The user can choose to output all of the created conformers.

Simplified Distance Geometry in Key3D We describe here the simplified distance geometry technique that has been used in generating 3D structures in Key3D. Good reviews of the standard distance geometry method have been published,[19,20] so we will focus on our modifications.

First, a distance bounds matrix that contains the maximum and minimum possible distances (upper and lower bounds) between each atom pair in the molecule is prepared as in the standard distance geometry method. The simplified method in Key3D skips the succeeding steps of smoothing, metrization and embedding, among which the metrization step requires a particularly large computational time. Instead of the crude Cartesian coordinates generated through these standard procedures, Key3D initially places atoms at random positions in a large 3D cube, 90 Å on a side. Then, minimization of an error function, which corresponds to the final step of the standard distance geometry, is performed. The error function is as follows:

$$E=\sum_{i=0}^{N-1}\sum_{j=i+1}^{N}\left(\max\left[0,\left(\frac{d_{ij}^2}{u_{ij}^2}-1\right)^2\right]+\max\left[0,\left(\frac{2l_{ij}^2}{l_{ij}^2+d_{ij}^2}-1\right)^2\right]\right)$$
$$+c\sum_{k=0}^{N_{\text{chiral}}}(V_k-V_k{}^*)^2 \qquad (1)$$

where N is the number of atoms, d_{ij} is the distance between atoms i and j, u_{ij} and l_{ij} are the upper and lower bounds, N_{chiral} is the number of chiral constraints, V_k is the chiral volume for a given chiral constraint k, and V_k^* is the desired value on that chiral volume. This error function is commonly used in the standard distance geometry calculation. The first term is for distance errors, and the second term corresponds to chiral errors. The coefficient c for the chiral term has been introduced so that Key3D calculation reaches more desirable structures in terms of both accuracy of bond lengths and angles and correctness of stereochemistry and planarity. The value of 0.25 has been given to the coefficient c for asymmetric centers and 1.0 for planar groups.

The stereochemistry is enforced to be the desired one by means of the chiral term of the error function, and the regiochemistry can also be controlled to some degree by the distance bounds of 1,4 atom pairs (*i.e.*, atoms separated by three bonds). Nevertheless, the minimization of the error function starting from random coordinates occasionally results in a 3D structure with incorrect stereochemistry or regiochemistry. To deal with this problem, the "structure confirmation loop" has been implemented. Key3D checks the structure after the structure optimization step, and if it has undesired stereochemistry or regiochemistry, Key3D goes back to the generation of random coordinates to try to construct another structure.

Evaluation of Key3D. Test Set A dataset of X-ray structures that Sadowski *et al.*[21] selected from the Cambridge Structural Database (CSD)[22] was used to evaluate the performance of Key3D. This test set does not focus on the drug-like structures that we place emphasis on, but it contains a wide variety of small molecules with high quality structures. The binding coordinates of ligand molecules are available from the Protein Data Bank (PDB),[23] but they were not used in our evaluation study. This is because the ligand coordinates in the PDB data have lower resolution than those in the CSD, and moreover, Key3D does not aim at the reproduction of bound conformation as stated in the introduction.

The test set contains a total of 639 structures, of which compounds including elements other than H, C, N, O, Si, P, S, F, Cl, Br and I are not targets of Key3D. We therefore used 598 molecules, excluding such compounds, in this validation study. The test set includes 213 molecules with one or more asymmetric centers (up to 17 asymmetric centers).

Sadowski *et al.* have carried out detailed comparative evaluation using this dataset for various 3D model builders including CORINA and CONCORD, and the results were reported in their paper.[21] However, their comparative study was carried out more than 10 years ago, so direct comparison of the performance between Key3D and the other methods studied in the paper is difficult.

Evaluation Tests Information only on atomic species, connectivity and stereochemistry was input to Key3D. Three cycles of conformer loop were designated in the Key3D calculation. The dielectric constant used in the MMFF structure optimization was set to 78.0.

We used the same test procedures as performed by Sadowski *et al.*,[21] with several modifications. The conversion rate, the number of program crashes, the number of asymmetric centers of incorrect stereochemistry, and the number of double bonds of incorrect regiochemistry were checked. To check CPU time, the calculation time on our Linux machine (Intel Xeon CPU 3.06 GHz) was measured. Close contact ratio (CCR),[21] which is the ratio be-

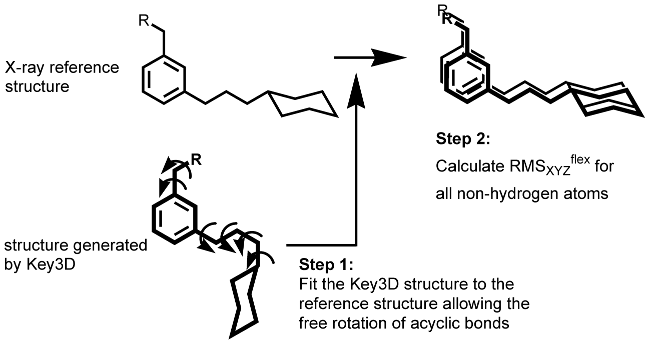

Fig. 2. Procedure of the Flexible-Fitting Test

tween the actual interatomic distance and a minimal acceptable distance, was measured as in the original procedure. The root-mean-square deviation (RMSD) for non-hydrogen atoms of the Key3D structure from the X-ray reference coordinates was calculated, as a measure of the accuracy of the created structures. We calculated RMSD for atoms included in ring systems ($\mathrm{RMS}_{\mathrm{XYZ}}{}^{\mathrm{rings}}$), by means of the procedures described in the literature.

We also checked the occurrence rate of convergence error in single-point SCF calculation of the Key3D structure by using the MOPAC93 program,[24] although this test was not included in the report by Sadowski *et al*. The MO calculation occasionally fails to converge even for a 3D structure that seems to have been converted without problems, due to insufficient accuracy or energetic instability of the structure. We used the occurrence rate of convergence error in the MOPAC calculation as a measure of the suitability of the generated structures for general computational chemistry. The MNDO method was used for single-point SCF calculation.

Flexible-Fitting Test In the evaluation by Sadowski *et al.*, the RMSD were calculated in terms of the entire structure ($\mathrm{RMS}_{\mathrm{XYZ}}$), the ring moieties ($\mathrm{RMS}_{\mathrm{XYZ}}{}^{\mathrm{rings}}$), and the torsion angles along the acyclic bonds ($\mathrm{RMS}_{\mathrm{TA}}{}^{\mathrm{chains}}$). However, many molecular design methods in current use (*e.g.*, automated flexible docking, molecular superposition) are able to take into account the rotation of acyclic bonds, so there is a decreasing need for input structures in which the conformation of the acyclic moiety coincides with that in the crystal structure. Therefore, $\mathrm{RMS}_{\mathrm{XYZ}}$ and $\mathrm{RMS}_{\mathrm{TA}}{}^{\mathrm{chains}}$ are considered to be less important, and were not included in this study. The factors that are more important for current molecular design tools are accurate geometries (bond lengths, bond angles, ring conformations), and accurate orientations of substituents in the ring systems (*i.e.*, axial or equatorial position) of the input structures. So, we have developed a procedure for calculating RMSD of all non-hydrogen atoms after fitting the generated structure to the reference crystal structure while allowing the rotation of acyclic bonds ($\mathrm{RMS}_{\mathrm{XYZ}}{}^{\mathrm{flex}}$). This test procedure is named the flexible-fitting test. Figure 2 summarizes the method. In this procedure, the RMSD is minimized with the torsion angles along acyclic bonds of the generated structure as variables, by using the Simplex method.[25] If $\mathrm{RMS}_{\mathrm{XYZ}}{}^{\mathrm{flex}}$ is close to 0 Å, the geometry of the rigid portion, *i.e.*, overall bond lengths and angles, ring conformations and the orientations of ring substituents, closely coincides with that of the reference structure.

Results and Discussion

Basic Performance Table 1 summarizes the basic performance of Key3D. Key3D was able to convert all of the 598 molecules used in this study (*i.e.*, conversion rate = 100%). This test set corresponds to 93.6% of the 639 test structures in the original set. Key3D experienced no program crash during the calculations, and its robustness has been demonstrated. As for the CPU time, Key3D required an average of 0.49 s per molecule on our Linux machine (Intel Xeon CPU 3.06 GHz). Key3D is slower than the rule-based methods, but it is fast for a method based on distance geometry, which is generally time-consuming. The calculation time of Key3D is acceptable even for handling large compound libraries. No erroneous stereochemistry or regiochemistry was found in the Key3D structures.

Table 1. Performance of Key3D

Evaluation item	
Conversion rate (%)	94 (100)[a]
Program crashes	0
CPU time(s) per molecule	0.49
Erroneous stereochemistry	0
CCR>0.8 (%)	100
Convergence error in MOPAC 1SCF calculation	0

a) Conversion rate of Key3D was 100% for the structures containing only H, C, N, O, Si, P, S and halogens, whereas it was 94% for all the original test set.

Table 2. Accuracy of the Ring Structures Created by Key3D

$\mathrm{RMS}_{\mathrm{XYZ}}{}^{\mathrm{rings}}$ (Å)	No. of structures (Percentage of the total)
<0.3	532 (89%)
0.3—0.6	39 (7%)
0.6—0.9	18 (3%)
0.9—1.2	7 (1%)
1.2—1.5	2
>1.5	0
Total	598

Sadowski *et al.* established a criterion for acceptable structures without intramolecular clashes, based on analysis of the CCR values of the original X-ray structures; the criterion is CCR>0.8. All of the 3D structures generated by Key3D satisfied this criterion. It is a major advantage of methods based on distance geometry that few close contacts are included in the created structures, because severe clashes between atoms can be resolved during the step of minimizing the error function. Moreover, Key3D contains the further refinement step of structure optimization based on the MMFF force field so as to relieve structural distortion and intramolecular clashes that cannot be resolved only by the simplified distance geometry step.

In the single-point SCF calculation by using MOPAC93, all of the Key3D structures were successfully calculated without convergence failure. Apart from this validation study, we have rarely encountered MOPAC convergence failure for regular organic compound structures converted by Key3D. These results support the suitability of the Key3D structures for computational chemistry.

RMSD from Reference Crystal Structure Table 2 shows the distribution of the RMSD of the ring atoms ($\mathrm{RMS}_{\mathrm{XYZ}}{}^{\mathrm{rings}}$) at intervals of 0.3 Å. This value reflects the accuracies of bond lengths, bond angles and conformations of the ring systems. In 89% of the structures, Key3D accurately reproduced the cyclic parts of crystal structures (RMSD<0.3 Å), and demonstrated high accuracy.

The RMSD of ring systems in the Key3D structures were rather large in some cases. There are several possible reasons for this. First, the number of iterations for conformer loops might not be sufficient. Since only three iterations were executed in this test calculation, the conformer resembling the crystal structure may not be produced, especially in the case of highly flexible ring systems. Key3D is able to perform conformational searches in which more conformers are generated and can output all conformers. For seven molecules with $\mathrm{RMS}_{\mathrm{XYZ}}{}^{\mathrm{rings}}$ of more than 1.0 Å (Fig. 3), we carried out

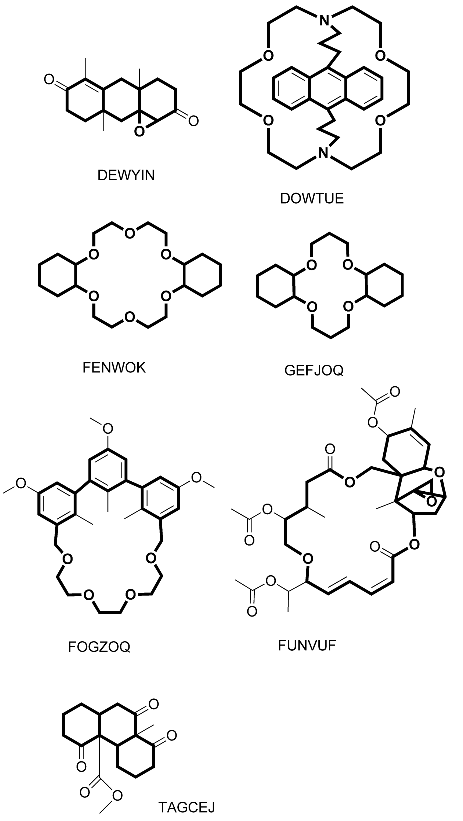

Fig. 3. Chemical Structures and CSD Refcodes of the Molecules with $\mathrm{RMS}_{\mathrm{XYZ}}^{\mathrm{rings}}$ of More Than 1.0 Å

The ring moiety in bold line was used in estimating $\mathrm{RMS}_{\mathrm{XYZ}}^{\mathrm{rings}}$ of each molecule. Stereochemical information is not shown for the sake of clarity.

the Key3D calculation setting the number of the conformer loop to 30. As is shown in Table 3, for all the molecules, a ring conformation closer to the crystal reference structure was found among the 30 generated conformers (the smallest RMSD: 0.05—0.69 Å). The effects of conformational search with increased iterations, taking the case of DEWYIN as an example, is shown in Fig. 4. The smallest-RMSD structure obtained by conformational search had a very similar ring conformation to the crystal structure (RMSD 0.13 Å; Fig. 4b), while the ring conformation generated through only three iterations of the conformer loop was significantly different from the reference one (RMSD 1.18 Å; Fig. 4a). The second possible reason for large RMSD of ring systems is that accurate estimation of the relative stability of generated conformers is still difficult. In the results of conformational search with 30 iteration loops, the ring conformations of the lowest-energy structures were totally different from those of the crystal reference structures in the cases of FENWOK and GEFJOQ (Table 3). The MMFF94 that we have adopted for energy calculation in Key3D is one of the most accurate and reliable force fields for small organic molecules. However, it should be noted that the effect of crystal packing influences the stability of the crystal structure. Since most of the MMFF parameters are based on high-level *ab initio* MO calculations in the gas phase, it may be difficult for the force field to reproduce the effects of crystal packing. We consider that the reproduction of this effect is not necessarily important for the purpose of molecular design.

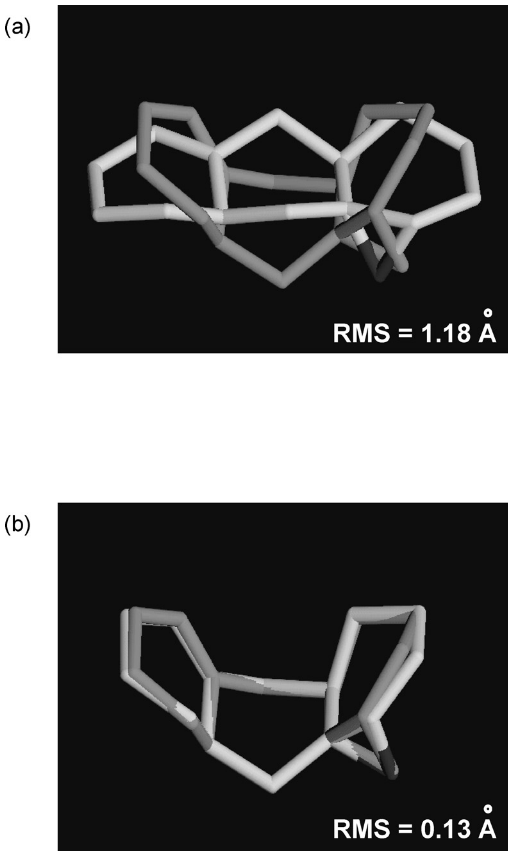

Fig. 4. Key3D Structures of DEWYIN, Obtained with Different Calculation Conditions; (a) Three Iterations of the Conformer Loop (Default Condition), (b) Conformational Search with 30 Conformer Loops

Each Key3D structure is compared with the reference crystal structure. The Key3D structure is shown in white, while the crystal structure is shown in gray. Only atoms included in the ring system are displayed for the sake of clarity.

Table 3. Effects of Conformational Search by Key3D on Accuracy of Ring Structures[a)]

CSD refcode	The 3D structure obtained by the default conversion condition	The lowest energy structure obtained by conformational search	The smallest-RMSD structure obtained by conformational search
DEWYIN	1.18 (53.0)	0.13 (49.7)	0.13 (49.7)
DOWTUE	1.42 (134.2)	0.83 (123.4)	0.52 (125.8)
FENWOK	1.21 (56.9)	1.21 (56.9)	0.69 (68.7)
FOGZOQ	1.16 (153.4)	0.97 (152.9)	0.50 (155.8)
FUNVUF	1.11 (151.0)	0.52 (131.6)	0.52 (131.6)
GEFJOQ	1.14 (55.2)	1.36 (46.7)	0.54 (61.8)
TAGCEJ	1.02 (65.8)	0.05 (59.5)	0.05 (59.5)

a) The RMSD of the ring atoms is shown in Å unit. The value in parenthesis shows the MMFF force field energy (kcal/mol).

Flexible-Fitting Test Figure 5 shows the results of the flexible-fitting test of Key3D structures. This test evaluates how accurately the generated structures can reproduce the corresponding crystal structures when free rotations of acyclic bonds are allowed, as in flexible docking and flexible molecular superposition. As shown in Fig. 5, 76% of the total Key3D structures coincided with the reference crystal structures with $\mathrm{RMS}_{\mathrm{XYZ}}^{\mathrm{flex}}$ of less than 0.3 Å. This result demonstrates high accuracy of the Key3D structures and the suitability of Key3D for use in current molecular design procedures. In our experience, an initial ligand structure with

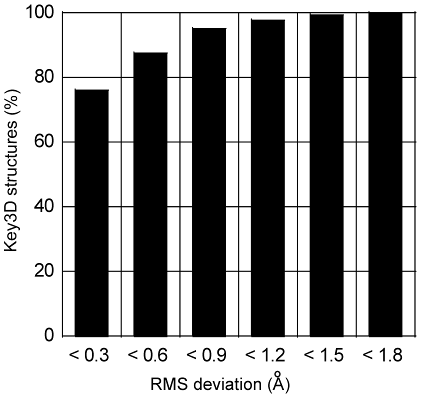

Fig. 5. Accuracy of the Key3D Structures Evaluated by the Flexible-Fitting Test

$\mathrm{RMS}_{XYZ}{}^{\mathrm{flex}}$ of less than 0.6 Å can afford an accurate docking model (in general, RMSD$<$1.2 Å). In this regard, 88% of the Key3D structures were of satisfactory quality. Moreover, further improvement of the accuracy is possible by increasing the number of cycles of the conformer loop. In this study, we used three cycles of the conformer loop. When we increased this to five cycles, the number of accurate Key3D structures ($\mathrm{RMS}_{XYZ}{}^{\mathrm{flex}}<0.3$ Å) increased by 2—3%, although more CPU time was required in proportion to the number of cycles of the conformer loop (data not shown).

In cases where rather large RMSD were observed in the flexible-fitting test, the difference in ring conformation was the major factor. Large RMSD were also observed when the orientation (axial or equatorial) of ring substituent(s) was inconsistent between reference and generated structures, no matter how accurate the other factors (bond length, bond angles, ring conformation) were. More elaborate conformational search by Key3D would lead to the acquisition of more appropriate structures in almost every case, as shown by the results for $\mathrm{RMS}_{XYZ}{}^{\mathrm{rings}}$.

Features of the Key3D Algorithm Here, we summarize the features of Key3D algorithm, based on the results of the evaluation tests described above. Key3D adopts a distance geometry method, which is more flexible than the rule-based approach that is inevitably restricted by the number and quality of internal rules and the library size. To improve the calculation speed of the distance geometry module, we have developed a highly simplified procedure centering on minimization of the error function. In consequence, both the accuracy of the generated structures and the calculation time are satisfactory for actual use in molecular design. The minimization step of the error function also contributes to the very low rate of conversion failure (in this evaluation study, there was no failure of conversion). Moreover, Key3D has been designed to provide good reproducibility of calculations, although a stochastic factor (*i.e.*, use of random numbers) is included in the algorithm. Another method of modified distance geometry depending on the refinement of error function has been developed by Spellmyer *et al.*[26] However, to our knowledge, their approach has not yet been validated as a practical 3D model builder. Key3D has the advantage of including all necessary functions for a practically useful and high-quality 3D builder, such as control of stereochemistry and structure optimization using a high-grade force field.

Key3D is also able to perform random conformational search. Since the conformational search by Key3D does not depend on a conformation library, which may restrict the variety of created conformers, Key3D can produce various stable conformations for any molecule. We showed above an example of the generation of a number of ring conformations. For the purpose of general 3D model building, the accuracy of Key3D structures can be improved by designating additional cycles of conformer loops. Thus, the user can easily adjust the balance between quality and speed in Key3D calculation.

Energy minimization of the intermediate structures constructed by the simplified distance geometry, on the basis of the MMFF force field, greatly contributes to the high quality of the Key3D final structures. In most 3D model builders, energy minimization is performed only roughly by using a simple force field, or no minimization is executed, because it requires significant computational time. However, we consider that elaborate structure optimization using an accurate force field is worth the cost of greater calculation time. Key3D has the advantage of directly generating reliable structures with no need for post-processing, and thus is suitable for practical use in actual drug design. For instance, when the Key3D structures were used for input structures in an evaluation test of flexible docking, high docking accuracy was achieved.[18] Furthermore, in our drug discovery projects based on virtual screening, far better results have been obtained by using Key3D than by using CONVERTER, which we had previously adopted.[27]

There remains some difficulty concerning the energy estimation. In our evaluation, the ring conformation in the lowest-energy structure generated by Key3D was not always consistent with that in the corresponding crystal structure. Nevertheless, ring conformers with small RMSD were found within a moderate energy range. Considering that more than one conformation may exist within the thermally accessible energy range, and that the most stable conformation may be different in the crystal, in solution, or in a protein-bound state, it is likely to be advantageous to provide several low-energy conformers as model structure candidates, for successful molecular design.

Acknowledgments We are grateful to Professor Johann Gasteiger and Dr. Christof Schwab (Universität Erlangen-Nürnberg) for kindly providing the coordinates of 639 CSD structures used in their evaluation scheme.

References and Notes

1) Present address: *Quantum Bioinformatics Team, Center for Promotion of Computational Science and Engineering, Japan Atomic Energy Agency; 8–1 Umemidai, Kizu-cho, Soraku-gun, Kyoto 619–0215, Japan.*
2) Sadowski J., Gasteiger J., *Chem. Rev.*, **93**, 2567—2581 (1993).
3) Gasteiger J., Rudolph C., Sadowski J., *Tetrahedron Comp. Methodol.*, **3**, 537—547 (1990).
4) Pearlman R. S., *Chem. Des. Autom. News*, **2**, 1—6 (1987).
5) Rusinko A., Sheridan R. P., Nilakantan R., Haraki K. S., Bauman N., Venkataraghavan R., *J. Chem. Inf. Comput. Sci.*, **29**, 251—255 (1989).
6) Crippen G. M., Havel T. F., *Acta Cryst.*, **A34**, 282—284 (1978).
7) Halgren T. A., *J. Comp. Chem.*, **17**, 490—519 (1996).
8) Halgren T. A., *J. Comp. Chem.*, **17**, 520—552 (1996).
9) Halgren T. A., *J. Comp. Chem.*, **17**, 553—586 (1996).
10) Halgren T. A., Nachbar R. B., *J. Comp. Chem.*, **17**, 587—615 (1996).
11) Halgren T. A., *J. Comp. Chem.*, **17**, 616—641 (1996).
12) Friesner R. A., Banks J. L., Murphy R. B., Halgren T. A., Klicic J. J., Mainz D. T., Repasky M. P., Knoll E. H., Shelley M., Perry J. K., Shaw

D. E., Francis P., Shenkin P. S., *J. Med. Chem.*, **47**, 1739—1749 (2004).

13) DesJarlais R. L., Sheridan R. P., Dixon J. S., Kuntz I. D., Venkataraghavan R., *J. Med. Chem.*, **29**, 2149—2153 (1986).

14) Leach A. R., Kuntz I. D., *J. Comp. Chem.*, **13**, 730—748 (1992).

15) Mizutani M. Y., Tomioka N., Itai A., *J. Mol. Biol.*, **243**, 310—326 (1994).

16) Jones G., Willett P., Glen R. C., Leach A. R., Taylor R., *J. Mol. Biol.*, **267**, 727—748 (1997).

17) Kato Y., Inoue A., Yamada M., Tomioka N., Itai A., *J. Comput.-Aided Mol. Des.*, **6**, 475—486 (1992).

18) Mizutani M. Y., Takamatsu Y., Ichinose T., Nakamura K., Itai A., *PROTEINS*, **63**, 878—891 (2006).

19) Crippen G. M., Havel T. F., "Distance Geometry and Molecular Conformation," ed. by Bawden D., Research Studies Press (Wiley), New York, 1988.

20) Blaney J. M., Dixon J. S., "Reviews in Computational Chemistry," Vol. 5, Chap. 6, ed. by Lipkowitz K. B., Boyd D. B., VCH. Publishers, New York, 1994, pp. 299—335.

21) Sadowski J., Gasteiger J., Klebe G., *J. Chem. Inf. Comput. Sci.*, **34**, 1000—1008 (1994).

22) Allen F. H., Davies J. E., Galloy J. J., Johnson O., Kennard O., Macrae C. F., Mitchell E. M., Mitchell G. F., Smith J. M., Watson D. G., *J. Chem. Inf. Comput. Sci.*, **31**, 187—204 (1991).

23) Berman H. M., Battistuz T., Bhat T. N., Bluhm W. F., Bourne P. E., Burkhardt K., Feng Z., Gilliland G. L., Iype L., Jain S., Fagan P., Marvin J., Padilla D., Ravichandran V., Schneider B., Thanki N., Weissig H., Westbrook J. D., Zardecki C., *Acta Cryst.*, **D58**, 899—907 (2002).

24) MOPAC93, Fujitsu Limited, Tokyo, 1993.

25) Nelder J. A., Mead R., *Computer J.*, **7**, 308—313 (1965).

26) Spellmeyer D. C., Wong A. K., Bower M. J., Blaney J. M., *J. Mol. Graph. Model.*, **15**, 18—36 (1997).

27) Mizutani M. Y., Itai A., *J. Med. Chem.*, **47**, 4818—4828 (2004).